A SURVEY OF PROOF THEORY II

G.KREISEL

Stanford University

Abstract

This paper explains recent work in proof theory from a neglected point of view. Proofs and their representations by formal derivations are treated as principal objects of study, not as mere tools for analyzing the consequence relation. Though the paper is principally expository it also contains some material not developed in the literature. In particular, adequacy conditions on criteria for the identity of proofs (in § 1c), and a reformulation of Gödel's second theorem in terms of the notion of canonical representation (in § 1d); the use of normalization, instead of normal form, theorems for a direct proof of closure under Church's rule of the theory of species [in § 2a(ii)] and the uselessness of bar recursive functionals for (functional) interpretations of systems containing Church's thesis [in § 2b(iii)]; the use of ordinal structures in a quantifier-free formulation of transfinite induction (in § 3); the irrelevance of axioms of choice to the explicit realizability of existential theorems both for classical and for Heyting's logical rules (in §4c) and some new uses of Heyting's rules for analyzing the indefinite cumulative hierarchy of sets (in §4d); a semantics for equational calculi suitable when terms are interpreted as rules for computation [in Appl. Ia(iii)], and, above all, an analysis of formalist semantics and its relation to realizability interpretations (in App. Ic). A less technical account of the present point of view is in [21].

Introduction

The main results reported in [18], which will be referred to as (SPT), concern *proof theory as a tool for studying logical consequence*; specifically, logical consequence from (subsets of) the usual axioms of analysis or set theory. The methods of proof theory are needed to establish the independence of ω-consequences such as various induction principles since, at least at present, we do not have *manageable* models which are non-standard with respect to the integers. Since consequence is a relation between formulae and sets of formulae while proof theory is concerned with derivations, the bulk of the proof theoretic machinery gets "lost" in the statements of theorems. Thus the ratio

interest of results/effort involved

is unsatisfactory. In the present complement to (SPT) I shall try to improve
this ratio by formulating results that do refer explicitly to the derivations and
not only to the consequence relation.

In SPT and particularly [20], I stressed another unsatisfactory aspect of
present day proof theory which is also connected with, so to speak, the shady
role of derivations; specifically, the lack of a clear explanation of the *choices*
of formal rules (studied or used). There I looked at the problem from the
point of view of Hilbert's programme or, more generally, of an analysis of
different *kinds* of intuitive proof; in short from an epistemological point of
view. Here I wish to emphasize formal results and problems concerning *rela-
tions* **between** *proofs*, for example the *identity relation between proofs de-
scribed by formal derivations of a given system* (and related topics, in § 1). It
must be expected, and it seems to turn out, that this study cuts across epis-
temological distinctions; specifically, that a formal theory of some of these re-
lations will apply uniformly to proofs of, epistemologically, obviously dif-
ferent kinds. This is of course quite consistent with the impression of mathe-
maticians, who are certainly occupied with the notion of identity of proofs,
that (epistemo) logical distinctions between different kinds of proof are not
useful to mathematical practice.

The formal results mentioned are mainly about *normalization procedures
for natural deductions* first developed systematically by Prawitz [32] and
since then also by Martin-Löf, particularly in the present volume. This whole
chunk of proof theory was left out of SPT simply because I did not realize its
significance. Here it is to be remarked that this part of proof theory is some-
times connected with Gentzen's ideas on the *meaning of the logical operations
as rules for their use,* about which I expressed reservations in SPT, bottom of
p. 329. I still have reservations which will be explained in App. I with some
help from the heuristically useful distinction between *computation* and
inference tules. But *modulo* certain conjectures in § 1, the formal relations
treated by Prawitz have significance independently of the possibly problem-
atic ideas that led to them.

§ 2 and § 3 formulate and expand some simple basic points which are
hidden in the Technical Notes of SPT, resp. in the very discursive discussion
of ordinal structures, pp. 333–339.

In § 4 I develop a bit the remark above to the effect that (familiar) logical
distinctions are not much used, by going over Heyting's own explanation of
the meaning of the logical operations and his formal laws. All this was *intended*
for constructive mathematics, to be interpreted in terms of constructive proofs

and functions. But what is actually asserted makes little use of any detailed knowledge of constructivity. Needless to say, this fact does not cast doubt on the interest of a closer analysis; on the contrary, I mention the fact to stress the need for discovering sharper requirements on such an analysis.

It goes without saying that a reader of this article will have studied Prawitz' contribution to the present volume where the main ideas of normalization are described. (After all, as mentioned already, SPT II was written principally because the subject of normalization had been neglected in SPT.) Here I surround these main results, in the style of my earlier expositions, with comments on their significance, some more or less obvious consequences, and open problems. Experience suggests that, at least at the present stage of proof theory, there is need for this kind of information to be sure that we are really dealing with "main" results. Actually the text below overlaps a little with Prawitz' [37] in that there are *some* indications of the main ideas; the indications are given not because I thought it necessary (or possible for me) to improve on his exposition of basic principles but because the two papers were written during the same period and so I could not refer to the details of [37].

1. Proof versus consequence

The general nature of our problem is quite clear. Consider *formal rules* which are intended to formalize certain *proofs*; in other words, we have syntactic objects, derivations, d which represent or decribe mental acts $\bar{d}$, the proofs (which carry conviction); cf. [22] p. 196 for an elementary discussion of the relation between d and $\bar{d}$, or the "mapping": $d \to \bar{d}$, which is a particular case of the general relation between words and the thoughts they express. Since we are dealing with a "small" class of words, we can hope for more precise results than are known for the more general (and more familiar!) relation. We must expect more detailed results for formal rules which are intended to codify reasoning about proofs, for example Heyting's systems, than for formal rules, specifically classical ones, whose intended interpretation refers to truth (and not, explicitly, to proofs); cf. §2a(i).

Given a property P of or a relation R between proofs our task is to find relations P_F, resp. R_F such that for all d of our formal system

$$P_F(d) \text{ iff } P(\bar{d}) \text{ and } R_F(d,d') \text{ iff } R(\bar{d},\bar{d}') .$$

For exposition, we shall reverse this procedure, and first describe some formal relations P_F, R_F which will then be used to state the facts about the objects of principal interest, namely properties and relations of proofs.

Superficially, the single most striking consequence of the view of proof theory just described, is certainly this. *Different styles of formalization previously judged only by aesthetic criteria of elegance or convenience acquire independent significance;* cf. end of (a) and (c) below.

(a) *Normal derivations and conversions.* Ever since Gentzen's original work there has been stress on special "normal" derivations, with different analyses of what is essential about these derivations (cf. middle of p. 329 of SPT). Here the special role of these normal derivations will be that they serve as *canonical representations* of all proofs represented in the system considered, the way the numerals are canonical notations for the natural numbers.

A minimum requirement is then that *any derivation can be normalized,* that is transformed into a unique normal form by a series of steps, so-called "conversions", each of which preserves the proof described by the derivation. This requirement has a formal and an informal part:

(α) The *formal* problem of establishing that the conversions terminate in a unique normal form (independent of the order in which they are applied).

(βi) The *informal* recognition (by inspection) that the conversion steps considered preserve identity, and the informal problem of showing that

(βii) distinct, that is incongruent normal derivations represent different proofs (in order to have unique, canonical, representations).

For examples of remarkable progress with the formal problem see the work of Martin-Löf and Prawitz in this volume. The particular conversion procedures considered evidently satisfy requirement (βi) since each conversion step merely contracts the introduction of a logical symbol immediately followed by its elimination. Such a contraction clearly does not change the proof described by the two formal derivations (before and after contraction).

Discussion. We can now restate the significance of the distinction described in SPT, p. 329, Ex. 2, between completeness of the normal derivations (in the sense that to each derivation there exists some normal derivation with the same end formula, for short: a normal form theorem) and normalization by explicitly prescribed, deterministic or non-deterministic procedures. *A normal form theorem leaves open the informal requirement (βi) above.* Provided of course proper attention is paid to (βi) (as in the remarks on the particular procedure of Prawitz and Martin-Löf), *a normalization theorem is the proper tool for the study of (βi).*

As stressed in SPT, p. 364, contrary to a common misconception the distinction is *not* to be analyzed in terms of constructivity, in the sense that model theoretic proofs of normal form theorems *can* be made constructive. SPT treats first order logic. In the case of higher order logic, if "constructive"

is interpreted as meaning: formalized in the theory of species, Prawitz' proof of the normal form theorem in [33] is made constructive in Technical Note II of [19]. (As is well-known, cf. SPT §12, pp. 351, 352, this is not the meaning of "constructive" used in the bulk of existing proof theory.) However, generally speaking, the natural proof of a normalization theorem has turned out to be constructive (in the sense described) as it stands[1].

Naturally the point of view, stressed at the beginning of this subsection concerning the significance of different styles of formalization, requires us to reconsider familiar criteria of "equivalence". Formalizations which we find significantly different from the present point of view, may have the *same set of theorems*, and this fact will in general be proved by quite elementary methods. A fortiori, they are equivalent as far as those subclasses of theorems are concerned which are involved in so-called proof theoretic strength. (This measure, proper to Hilbert's programme, is quite inappropriate here.) As a corollary the point of view opens up new areas of formal work. Consider specifically a calculus of sequents and a system of natural deduction. Inspection shows that many cut elimination procedures for calculi of sequents do not obviously satisfy the informal requirement (βi), and also that the normalization procedure for systems of natural deduction does not correspond to a particularly natural cut elimination procedure. Consequently we shall have to look at the usual formulations of a calculus of sequents, vary the *order* in which different cuts are considered, and see if different orders lead to non-congruent cut-free derivations (or even, for some systems, that the procedure does not terminate at all). In this case the formal requirement (α) would not be fulfilled (in contrast to the case of natural deduction treated by Prawitz and Martin-Löf). Note that (α) follows from (β) since by (βi) all conversions

---

[1] It often happens when we ask for *more* (information), the *natural* solution is more elementary; for an analysis of Hilbert's ϵ-substitution method from this point of view see [17] p. 168, 3.351. For actual research it would be important tr be more specific in order to know which methods are likely to succeed (naturally) with a given kind of problem! But it is hard to analyze the existing evidence. Certainly, "theoretical" results on the existence of a logically more elementary solution are available, but they do not seem relevant to the practical question of discovery. For example, let A be a first order assertion about fields; if A is true for the field Q of real rationals the proof may require transcendental axioms; if A holds for all real fields, not only for Q, there *exists* a first order proof (from the first order axioms for real fields). But the most natural proof may well involve definitions of a real closure and set theoretic operations on it which are not even definable in the (first order) language of fields. Similarly the mere existence of a constructive proof of the normal form theorems does not ensure that the natural, or at least first, proof *is* constructive. (In the first order case it was, in the second order it was not.)

are supposed to preserve identity of proof and by (βii) *must* terminate in congruent normal forms; but (α) alone does not ensure (β). For more detail on these matters, see p. 91 of Prawitz' monograph [32].

The kind of work proposed here is the sort of *Kleinarbeit* which is generally needed to support a genuine hypothesis (here: concerning the *adequacy* of the normalization theorem for identity criteria on proofs) as opposed to a mere mathematical fancy. Specifically it is needed because we can hardly hope that existing formalizations such as those in [32] are *exactly* right for the new applications, for instance those of (b) and (c) below. After all, the systems were developed for other reasons, logical or aesthetic.

Remark. The principal technical tool used by Prawitz and Martin-Löf for normalizing deductions in the theory of species was introduced by Girard [56]. Independently Girard used his ideas to establish the termination of a certain cut-elimination procedure, thus giving a new proof of Takeuti's "fundamental conjecture" for analysis. I have not had an opportunity to study whether Girard's procedure preserves identity of proofs.

(b) *Normal derivations of existential formulae.* Evidently, besides relations on proofs it makes good sense to consider relations between *proofs, assertions* and *definitions*; correspondingly we have formal relations between *derivations, formulae* and *terms*. There is a very natural relation which has been much discussed in the (constructive) literature, namely

the definition $\bar{t}$ is provided by the proof $\bar{d}$ of the existential assertion $\overline{\exists xA}$;

cf. Problem 1 on p. 125 of [19] for an elementary discussion. Though quite special it will turn out that this relation is quite useful.

An important property of *normal derivations* (as defined by Prawitz et al.) is that they allow us *to read off t from a derivation d of $\exists xA$*; more precisely there is a mechanical method of obtaining the term t from d (which term we recognize as defining the realization provided by $\bar{d}$). Obviously there is a mechanical method of deciding whether d is a derivation of an existential formula. For reference below note that t may contain parameters not occurring in $\exists xA$, for instance if the endpart of d derives $\exists xA$ from $\forall xA$ or from $A[x/t_1] \vee A[x/t_2]$.

Note that, by (a), the analysis of the relation above requires genuinely a normalization and not only a normal form theorem.

(c) *Formulating one-sided adequacy conditions on criteria for the identity of proofs.* As stressed by Prawitz [37], his normalization procedures obviously

preserve identity of proofs. Our problem here, (βii) in (a), of deciding whether convertibility to congruent normal forms is *equivalent* to identity of proof, is more delicate. I shall formulate a, demonstrably only partial criterion[2] and afterwards go into the general nature of the problem. To repeat from the discussion in (a): it is not claimed that *exactly* the conversion relations studied in the literature are adequate. (*Added in proof*: see also fn. 20.)

We consider deductions in predicate logic, not arbitrary ones, but of formulae of the form

$$B \to \exists x A .$$

(i) We suppose that the relation $\equiv$ between formal derivations (with the same end formula) evidently preserves identity of proofs, i.e., in symbols

$$d_1 \equiv d_2 \Rightarrow \bar{d}_1 = \bar{d}_2 .$$

(ii) For further progress we try to derive from $\bar{d}_1 = \bar{d}_2$, the primitive relation under study, some mathematically manageable consequence, say $M(d_1, d_2)$, fulfilling the purely mathematical condition:

(iii) $M(d_1, d_2) \Rightarrow d_1 \equiv d_2 .$

If we succeed we shall have "trapped" our primitive relation and can infer

$$d_1 \equiv d_2 \Leftrightarrow \bar{d}_1 = \bar{d}_2 .$$

A candidate for $M(d_1, d_2)$. Let d_1 and d_2 be derivations in predicate logic of $B \to \exists x A$. Consider any formal system F such as the theory of species which includes predicate logic and is normalizable, the normalization steps satisfying (βi) of (a); we do *not* need here that F satisfies (βii) (fortunately, since (βii) for predicate logic is the principal issue here!). For each substitution instance $B^* \to \exists x A^*$ of $B \to \exists x A$, let d_1^* and d_2^* denote the corresponding substitutions in d_1 and d_2 resp. For any derivation d^* of B^*, let d_1^0 and d_2^0 be obtained by joining d_1^*, resp. d_2^* to d^*. Both d_1^0 and d_2^0 are derivations of $\exists x A^*$ in F and, by (b), provide terms t_1 and t_2 as realizations.

$M(d_1, d_2)$ asserts: for all derivations d^* of substitution instances B^* the corresponding terms t_1 and t_2 define *extensionally* equal functions. Then

---

[2] Prawitz has helped me find the proof that the criterion is partial.

$M(d_1,d_2)$ satisfies (ii), extensional equality being manageable (cf. §2b below); hence $\neg M(d_1,d_2) \Rightarrow \neg \bar{d}_1 = \bar{d}_2$.

But $M(d_1,d_2)$ does *not* generally satisfy (iii): it provides only a *partial* criterion that is, there are d_1 and d_2 for which $d_1 \neq d_2$, but the use of M described above does not ensure that $\bar{d}_1 \neq \bar{d}_2$ (in other words, at the present stage the exact significance of the relation $\equiv$ is open for such d_1, d_2).

Let B be true (and hence B^* derivable) and take for A a suitable formula $\exists x(x=c \wedge P)$ with P independent of x, e.g.,

$$\exists x(x=c \wedge [(p \wedge \neg p) \rightarrow (p \rightarrow p)]) .$$

Clearly *any* proof will provide the realization c. But we have two clearly different proofs of

$$(p \wedge \neg p) \rightarrow (p \rightarrow p)$$

(and hence of the whole formula); namely one ignores the conclusion (ex falso quodlibet), the other ignores the premise ($p \rightarrow p$ being valid).

Discussion. To get perspective on the proposed (partial) criterion, it seems instructive to consider the familiar analysis of the notion of *logical validity* (from Frege's formulation of rules of inference to Gödel's completeness proof). The primitive notion of identity of proof is to correspond to logical validity, perhaps with the subclass of deductions considered here corresponding to the restriction to *first* order language. Then (i) above corresponds to the fact that Frege's formal rules evidently preserve logical validity. The "manageable" property M in (ii) corresponds to the mathematical property of validity in all countable domains (Skolem-Löwenheim). And the parallel to (iii) is provided by Gödel's completeness proof. For *subclasses* of the first order language, for instance for purely universal formulae, we can replace "validity in countable domains" by "validity in finite domains". For second order formulae we do not have formal rules (as is well-known), but also we do not have an equally manageable property M, second order validity being sensitive to open questions about the existence of large cardinals (cf. [22], pp. 190–191). In the latter case we operate with partial criteria, obtained by reflection on specific matters such as the existence of, say Π_1^1-indescribable cardinals. It is safe to say that proposed adequacy conditions for identity criteria need not be complete (but of course they must settle some interesting open question).

A much more serious objection to the criterion (than its partial character)

is this. The proposed proofs of adequacy *evade* rather than solve questions about the nature of the identity of proofs or, indeed, about the nature of proofs. For what we propose to show, for specific derivations, is that questions of the adequacy of $\equiv$ can be settled without closer analysis of the concepts involved. But, at the present time, the criterion above has a pedagogic use: it corrects the common assumption that "nothing" precise (and reasonable) can be done on questions about synonymity of proofs.

Remark. At the risk of trying to explain *obscurum per obscurius,* I should point out a striking analogy between the problem of finding (α) a conversion relation corresponding to identity of proofs and (β) an "equivalence" relation such as that of *combinatorial equivalence* in topology corresponding to a basic invariant in our geometric concepts. Quite often it is difficult to formulate explicit adequacy conditions on (β) in advance (our confidence in a proposed equivalence relation may well depend on the quality of the proposer). The use of different styles of formalization mentioned in (a) above, may be compared to the use of different coordinate systems in geometry, where one system is often particularly suited for the study of a specific geometric relation, "suited" not only in the sense of being manageable, but the passage from one co-ordinate system to another may introduce singularities without geometric meaning such as the indeterminateness of the second polar co-ordinate at the origin.

(d) *Further illustrations: Gödel's second theorem.* The reader who shares my confidence in striking examples (and perhaps also my skepticism about isolating *the* central problem at an early stage of research) may wish to look at familiar material which is most naturally formulated in terms of *proofs* rather than *consequence.* I think Gödel's second theorem is an excellent example.

We shall consider *consistent* formal systems. So, as far as the set of theorems is concerned, the question whether A and $\neg A$ both occur, will not arise at all!

Here a moment's thought shows that, in vivid language, *it is not even sufficient to identify a formal system with its set of deductions, we have to consider the manner in which we verify that a syntactic object is a deduction,* in short we have to consider the formal rules. (I shall take these rules to be given by their production rules.) For, consider any formal system F you know, and define F^* as follows (cf. [17], p. 154, 3.221).

An object d is a deduction of F^* if it is (i) a deduction of F and (ii) for all deductions d', d'' of F (with Godel number) $\leq d$, d' does not lead to a formula which is the formal negation of the formula derived by d''.

Now, if F is consistent, F^* has *exactly* the same deductions as F, only the verification of the property of being a deduction is longer, corresponding to step (ii). Evidently there is a perfectly elementary proof of the consistency of F^*. Equally evidently, different (in fact, *not* demonstrably equivalent) formulae *express* the proof relations of F and F^* and hence their consistency. The formal notion of *canonical representation* in 3.222 on p. 154 of [17] gives an *analysis of the concept of* **expressing** *the proof relations of F and F* which is adequate for the present problem* (of formulating Gödel's second theorem properly). Though not stated in [17], similar conditions determine, up to provable isomorphism, the *canonical* Gödel numbering of finite sequences (of $a_1, ..., a_n$ say) with the following structure: the empty sequence is a distinguished element, and for each a_i, adjunction of a_i and its inverse is an operation; cf. e.g. [23], p. 256, 2.5.3.

Remark. There is an additional point to be observed here which, however, does not involve attention to derivations; consequence is sufficient. It is sometimes said that Gödel's theorem applies to systems which "contain arithmetic" (or, more explicitly, *modus ponens* and are complete for numerical arithmetic). Clearly, in the ordinary sense of the word, F^* contains as much arithmetic as F! The natural formulation runs quite simply as follows.

Given F and formulae A_1, A_2, A_3 of F, one of the following cannot be derived in F:

A_1 expresses (in terms of canonical representations) that F is closed under modus ponens, and A_1 holds.

A_2 expresses that F is complete for numerical arithmetic (in fact complete with respect to a specific primitive recursive predicate, of course canonically represented), and A_2 holds.

A_3 expresses that F is consistent, and A_3 holds.

It is perfectly natural that in Gödel's original paper not much attention was given to a proper formulation. First of all, at the time the importance of normal derivations had not been realized, and hence, for the usual systems, formulae A_1 and A_2 could always be found. Secondly, as far as Hilbert's programme was concerned the absence of A_1 or A_2 constitutes the *same kind of inadequacy* as of A_3: we should have no reason to suppose that F codifies mathematical practice. (I say "the same kind" because of course Hilbert's programme, as he formulated it, does not require the adequacy conditions to be proved within the system studied.)

Summarizing, we see that not only deductions, treated as extensional objects, are relevant here (over and above the set of consequences), but even additional information or "structure", namely the sequence of operations involved in building up the deductions.

(e) *Note on the consistency problem* (supplementing SPT, p. 323 and pp. 360–361). The following points are elementary.

The consistency *results* do not express at all well the mathematical content of the consistency proofs; specifically (by SPT, Technical Note II, on reflection principles, conservative extensions and commuting models) we now have the concepts needed to express their content both elegantly and much more informatively; cf. also the end of b(ii) in App. I. Also (cf. SPT, Technical Note I) some of Hilbert's assumptions connecting consistency proofs with *reliability* seem unjustified. But, to avoid misunderstanding, it should be remembered that we *do* have perfectly genuine consistency proofs in the sense that the metamathematical methods used are *patently* more elementary than the intended interpretation (which led to the formal system studied; and not only in the case where there are simply doubts about the consistency as is SPT, p. 325, 1. 12–13). For example, in the case of arithmetic with induction restricted to purely universal formulae, the methods used by Herbrand or Gentzen are evidently more elementary than the set theoretic conception of the structure of arithmetic; equally, we have of course also model theoretic methods, where an ingenious model may use significantly more elementary existential assumptions than the obvious or "standard" model; cf. SPT, Technical Note IV. These observations are quite consistent with the fact that, *in special cases,* for example in Gentzen's consistency proof for arithmetic by ϵ_0-induction, the situation is more delicate. Since, practically speaking, the non-constructive conception of natural number is not genuinely problematic, and also the validity of ϵ_0-induction is certainly not immediately *seen*, the epistemological value of Gentzen's proof depends on a more delicate analysis of kinds of evidence. (Naturally more is written on such open questions which require attention than on the many consistency proofs with a clear-cut epistemological conclusion.) At least at the present stage of analysis the methods developed for his proof seem to me to have led to more interesting results for other application such as those discussed in (a) above than for the original consistency problem.

Remark. The fact that simple-minded problems concerning consistency or, more generally, *reliability* of principles should have led to significant distinc-

tions *among* reliable principles seems to me typical of a quite general pattern (just as in physics naive problems concerning *reality*, e.g. of light, led to significant distinctions *among* the objects of the real world). Naturally it is not suggested that the uses of these distinctions, specifically between normal and other derivations, have been exhausted! To mention only one example, implicit in (b) above, in connection with *length* or "feasibility" of constructions: the length of a non-normal derivation of $\exists xA$ may be much smaller than the numerical value of the realization which it provides. In short, "*ordinary*" *logical inference spoils feasibility*. (As observed by Takeuti, for normal derivations in his "cut-free" analysis this is not the case.)

2. Operations on derivations: syntactic transformations and functional interpretations

In contrast to §1 which treats matters that were neglected in SPT, the present section is concerned with a familiar distinction (cf. [17], pp. 159–160, 3.31) between the two general methods of proof theory mentioned in its heading. Typical syntactic transformations are normalization and cut-elimination procedures introduced by Gentzen. Typical functional interpretations are the no-counterexample-interpretation [14] or Gödel's [7]; also Herbrand's theorem can be viewed as such an interpretation (cf. [14]) though he himself did not do so (other points of difference are explained at the end of this section). Though implicit in the literature there are some simple general points about *areas of useful application* of these two methods which do not seem to have been stated explicitly. Since we are concerned with uses, the differences are more important than the similarities: given a problem we want a hint on which method is more likely to apply. Similarities will be gone into in App. I in connection with Prawitz' homomorphism between deductions and terms (proofs and functions-as-rules). For expository reasons we shall also defer, to §3, all *quantitative* refinements of these two methods in terms of ordinals (or, more pedantically, ordinal structures); there has been so much work on this aspect that a report on it here would distort the general picture.

Let me state quite briefly three practical conclusions which are formulated more precisely below.

(α) Syntactic transformations concentrate on *proofs* and need only very elementary functions; the interpretations (mentioned above, not realizability which is a hybrid) concentrate on *functions*, that is definition principles, and use only very elementary proofs, for detail, see (aiii) and (biii) below.

(β) Syntactic transformations are specially adapted for obtaining *derived rules*, generally speaking because normal derivations are so simple. Thus normal form theorems are often sufficient for this purpose.

(γ) Interpretations are useful for obtaining independence results from a *schema*. By varying the interpretation principle or the class of functionals considered, a whole schema can be made valid for an interpretation.

(a) *Syntactic transformations* map derivations $\epsilon\mathcal{D}$ into the (sub) class $\mathcal{D}_N$ of *normal* derivations. Naturally the applications may depend merely on structural properties of normal derivations, not on the particular transformation used, in which case a *normal form theorem* is sufficient. Obviously, if a normal form theorem is true for the pair $(\mathcal{D}, \mathcal{D}_N)$ there is always *some* recursive mapping: $\mathcal{D} \to \mathcal{D}_N$ preserving end formulae: given a derivation $de\mathcal{D}$, run through $\mathcal{D}_N$ until you find $d'\epsilon\mathcal{D}_N$ with the same end formula as d.

(i) In SPT, p. 329 and particularly p. 348(b) I stressed those normal derivations which possess the *subformula* property. The most striking consequence is that *if* an application of a normal form theorem depends on this particular property, for systems including induction it is an advantage to consider *infinite* normal proof figures[3]. (And even here we have limitations for recursive figures since there are, demonstrably, none satisfying the subformula property if bar induction is included; cf. [17], p. 167, 3.343 and SPT, p. 348(b).) At the present time, as far as I know, applications of normal forms in *classical* systems depend on the subformula property.

Digression on infinite proof figures. As a result of correspondence with Prawitz it seems desirable to expand my earlier remarks on this subject, in SPT, footnote 3 on p. 324 or p. 332(b). Firstly we have the point of principle, the *representation* of our thoughts, that is of proofs, by means of such infinite proof figures. There is no question that the latter are more explicit than finite proof figures, yet manageable enough. The (extensional) proof figures are *not complete* representations; as a minimum one would add a *description* (e.g. at

---

[3] We say proof *figure* and not proof *tree* because, for comparison with (ii), it is better to allow a formula A as a premise of two or more inferences, we do not require separate proofs for each occurrence of A. (As usual, we suppose the proof figure to be given not merely by a partial ordering relation, but with additional structure associating the premises of a formula to it.)

each node a description of the figure "below" the node) together with a proof
that the figure is properly built up. This corresponds to a familiar procedure
in technical proof theory spelt out in the clearest possible terms in [17] ,
p. 163 bottom; cf. also c(ii) below. From the general point of view stressed in
the present article concerning the problematic character of constructive logi-
cal operators, it is to be stressed that the proofs to be added to the extensional
proof figure (for a more complete representation of our intensions) establish
logic free assertions. At the time when SPT was written there was not even an
attempt in the literature of making explicit a *meaning* of finite derivations.
Only in the present volume do we have Prawitz's analysis [37] by use of his
validity predicate for deductions, albeit in a logically complicated form.
(Earlier uses of, formally similar, computability predicates made *no* contri-
bution to the present problem of explaining *what* is being coded by finite
derivations.) Secondly we have the mathematical point, in my opinion of
paramount importance for the intelligibility of the subject, of giving an *in-
trinsic meaning to the ordinals used in this part of proof theory*, mentioned in
SPT, p. 332 (b). The infinite proof figures were the *only* objects which had an
intrinsic connection with ordinals. The situation has changed with normaliza-
tion problems where *specific* well founded reduction figures are principal ob-
jects of study; cf. §3a(ii). Evidently the interest of these figures is contingent
on the interest of the *specific* reduction processes, and it was for that reason
that I went out on a limb in §1 to search for such an interest. (Here ends the
digression.)

   Returning now to applications of normal forms, for Heyting's systems (and
in contrast to the classical case considered earlier on) the subformula property
is not essential. Specifically, for normal forms *derived rules* are evident, gen-
erally of the following form: for A of suitable syntactic structure

   if $A \to \exists x B$ is derivable so is $A \to B[x/t]$ for some term t

and thus $\exists x(A \to B)$ where x is not free in A; cf. [17] , p. 160, 3.322. (They are
genuine derived *rules* in the sense that the implication $(A \to \exists x B) \to \exists x(A \to B)$
is not derivable.) Despite massive work in this area it is not generally realized
that such *conditional explicit definability results are much more useful than*
(the more familiar) *absolute ones*; for the simple reason that the *former ex-
tend automatically when axioms of the structure of A are added to the sys-
tem considered*; cf. [17] loc. cit. or p. 269 of [33] . As mentioned many
derived rules have the form above, including Markov's rule, discussed at
length in Troelstra's article in the present volume [46] . (It should perhaps be
remarked that, in formal arithmetic, $\forall x(A \lor \neg A) \to \neg\neg \exists x A$ may be derivable,
even if $\forall x(A \lor \neg A) \to \exists x A$ is not; take $T(e,e,x) \lor \forall y \neg T(e,e,y)$ for A with
parameter e; or constant e if $\exists x T(e,e,x)$ is not formally decidable.)

Two remarks seem to me pertinent here. Firstly, about the derived rules specifically, it should be noted that they are *not* linked to the constructive interpretation of Heyting's systems. On the one hand, as will be explained in §4, *not* even the disjunctive property is *required by* a sound interpretation; on the other as seen by adding (constructively) invalid axioms of the structure of A, constructive soundness is *not needed for* the derived rules. Secondly, about normal forms generally: though for some applications the subformula property is not required we do not expect (and do not get) all important properties of normal deductions in the first order case where we do have the subformula property. For example, the *interpolation lemma* can fail. (By SPT, p. 357 (iv) this does not cast doubt on the notion of normal form, but rather on the choice of language.)

(ii) An application which, at present, involves a *normalization* theorem in an interesting way, is the proof of: *closure under Church's rule*. The normal form theorem yields immediately the result: if $\forall x \exists y A(x,y)$ is derivable there is a recursion equation with number e_0 such that, for each numeral n, $A(n, \{e_0\}(n))$ is derivable where $\{e_0\}(n)$ denotes the numerical value of the function defined by e_0 at the argument n. The issue is to show that the universal formula $\forall x A(x, \{e_0\}(x))$ is also derivable in the system considered. (To be precise; I mean here the e_0 corresponding to the mechanical rule: run through the formal derivations till you hit the first one whose end formula has the form $A(n,m)$ with numeral m and associate this m to n.) Note that strictly speaking, we do not need a normalization theorem, but only this: given a proof of $\forall x \exists y A(x,y)$ in S we need a proof in S itself *that*, for each numeral n, $\exists y A(n,y)$ has a normal derivation in a suitable subsystem S_0; "suitable" means that the reflection principle for S_0 can be established in S. In Note II of [19], closure under Church's rule is obtained by combining Prawitz' *classical* proof [33] of the normal form theorem for the theory of species and a corollary to Spector's description [43] in his system Σ_4 of the provably recursive functions of classical analysis namely:

Classical analysis, with or without axioms of choice ([19], p. 135, 1.17-19), is *conservative* for $\forall E$ formulae over the theory of species.

In [19] the corollary is obtained from a *model* (in the theory of species) for Σ_4, which can now be replaced by a (simpler) proof of computability of Σ_4 by use of Girard's work [56] (or directly by [56] without mention of Σ_4).

As mentioned in §1a, the recent normalization theorems make this detour via [33] unnecessary. More importantly, by §1b they yield more; namely for a derivation d_n of $\exists y A(n,y)$, a term t_n defining the object $\bar{x}$ provided by the

proof $\bar{d}_n$. Since from a given derivation d of $\forall x \exists y A(x,y)$ we obtain d_n canonically by substitution, the normalization procedure also provides a recursion equation e_1 such that

$$\forall x A(x, \{e_1\}(x)) \text{ is derivable } and \ t_n^* = \{e_1\}(n)$$

(where t_n^* is some closed term obtained from t_n if the latter contains parameters; cf. [19] §1). Note that in general $\{e_0\}(n) \neq \{e_1\}(n)$, again by [19] §1; however e_1 cannot be said to be *the* rule provided by $\bar{d}$ as the intended meaning of the quantifier-combination $\forall \exists$ in $\forall x \exists y A$ [cf. App. Ia(i) for the defects of Kleene's T predicate for the present purpose].

(iii) Finally, a word about property (α) mentioned at the beginning of this section. The description of the (infinite) proof figures in (i) and the *immediate* reduction or normalization relation (not its transitive closure) in (ii) can be given by use of Kalmar elementary definitions; not even all primitive recursive ones are needed. This fact is plausible because well-orderings are involved and all recursive well-orderings are embeddable in an order preserving way in *Kalmar elementary ones* [12]; what has to be checked is that the embeddings are elementary too (and can be proved to be so by applying induction to elementary predicates; concerning this last and absolutely basic restriction, cf. [17] p. 165, 3.3322). It is in this sense that *syntactic transformations concentrate on proofs* (of elementary free variable statements) *and do not involve* (more than a minimum of) *functions*.

Though perhaps no more than a remark, the property of syntactic transformations just mentioned provides a useful concrete background for the discussion in §4 of Heyting's meaning of the logical operations[4].

---

[4] I have often stressed an analogy between set-theoretic and constructive foundations: while *types* (of sets) and *proofs* (at least of identities) are essential to foundations, they do not enter as objects of study into practice. Amusingly, we also have parallel objections! J.P.Serre once complained in conversation about the fuss logicians make about types since nobody uses the axiom of foundation anyway; and, he added, if you want it you simply *define* the collection of those sets which are hereditarily well founded. Scott, in [50], p. 239, 1.-10 to 1.-9, expresses very clearly similar misgivings about the role of proofs in constructive foundations. Pushed beyond reason, Serre's view blocks any chance of a convincing reason for Zermelo's "restriction" of Frege's inconsistent version of the comprehension principle and Scott's view blocks any chance, at least at present, of a non-circular explanation of implication; cf. also App. Ic(ii).

(b) *Functional interpretations* have the following general form ([17] p. 158, 3.301-3.303 or SPT p. 378–379). We associate with each formula A a formula $\exists s A_0(s)$ such that, if A is true so is $\exists s A_0(s)$ with s ranging over a "large" class F^+ of functions or functionals, and if $\exists s A_0(s)$ is proved by restricted methods, $\exists s A_0(s)$ is true if s ranges over a smaller, usually explicitly generated class F^-. If A_0 is logically complicated, the two statements are not comparable (because the ranges of the quantifiers inside A_0 will generally also be different). But if, for example, A_0 is purely universal, and if the smaller range of s is dense in the large range, for some topology on which the functionals considered are continuous, then $(\exists s \in F^-)A_0^-$ is stronger than $(\exists s \in F^+)A_0^+$. (This condition is satisfied by the interpretations mentioned at the beginning of this section.) In other words, such interpretations show clearly *what more we know when we have proved a theorem than if we only know that it is true.*

Though connections between syntactic transformations and functional interpretations do exist (and will be referred to repeatedly in this text), at the present time the single most important feature of the latter is their *simplicity*, particularly when classes of extensional functionals suffice for one's purpose, for example, for *independence results*. Now, if A is proved we not only have that $(\exists s \in F^-)A_0^-$ is true, but we have a mapping: $d \to s_d$, such that if d derives A in the formal system studied then $A_0(s_d)$ can be derived in the system, say $\mathcal{I}$, which supplies the interpretation. So, for independence of A, it is enough to find *some* class C of functionals satisfying $\mathcal{I}$, such that $A_0(s)$ is *false for all* $s \in C$. If A is independent for, so to speak, brutal reasons, any class C will do provided it satisfies very simple conditions; cf. also (iii) below.

It is not necessary here to go into the use of interpretations for independence results from schemata, i.e., property (γ) mentioned at the beginning of the section, because it is very well illustrated by Troelstra's exposition in [46]. (A reader who wants to compare independence proofs from a schema by means of normal forms with uses of interpretations may consider the independence of the schema M from IP in 5.2 of Troelstra's paper [46].)

(i) The well-known independence proofs of the law of the excluded middle, with variables f for sequences of type $0 \to 0$,

$$\forall f \exists n [fn=0 \lor \forall m(fm \neq 0)]$$

interpreted by

$$\exists N \forall f [f(Nf)=0 \lor \forall m(fm \neq 0)] \ ,$$

use only continuity of N for the product topology on sequences of natural numbers.

(ii) Again, often only the *recursive* character (in any of the various senses described in Troelstra's article [46]) of the functionals considered need be used; despite the fact that for any formalized $\mathcal{G}$, we shall have realizations (models) by recursively enumerable *subclasses* of the class of recursive functionals.

Of course the same applies to closure under Church's rule discussed in (a) (ii) above: the e's can be found in the recursively enumerable set of *demonstrable* recursion equations, that is equations that can be proved in the system considered to define a (total) function.

(iii) Coming now to *limitations* of functional interpretations we distinguish between those connected with the syntactic form of the interpretation $\exists s A_0$ of the formula A and those connected with the particular range of the variable occurring in A_0. As an example of the former we have formulae A where s does not occur in A_0 at all. This will happen, generally speaking when A is a purely universal formula (and quantifier-free formulae are decidable). In this case, the "brutal" method sketched cannot be expected to help in establishing the independence of A.

An interesting example of the second kind of limitation is provided by the socalled bar recursive functions of finite type introduced by Spector [43]. They demonstrably *satisfy the interpretation* [7] *of the negation of Church's thesis*. Indeed, by [43], p. 19, (12.1.1) they satisfy the interpretation of

$$\neg\neg\forall x \exists y \forall z [T(x,x,y) \vee \neg T(x,x,z)]$$

for Kleene's T, but also

$$\neg \exists e \forall x \exists v \forall z \{T(e,x,v) \wedge [T(x,x,Uv) \vee \neg T(x,x,z)]\}$$

since this formula is a theorem of first order intuitionistic arithmetic. But this conflicts with Church's thesis (indeed with its double negation)

$$\forall x \exists y A(x,y) \rightarrow \exists e \forall x \exists v [T(e,x,v) \wedge A(x,Uv)] \ .$$

Recently Professor Gödel has pointed out to me a more elegant proof. The system Σ_2, described [43] p. 6 is *formally inconsistent* with Church's thesis; (added in proof) cf. also chapter IX of [60]. Since T is decidable

$$\neg\neg\exists y\forall z\,[T(x,x,y)\vee\neg T(x,x,z)]$$

is a theorem of arithmetic. But so by axiom F

$$\forall x\neg\neg P\rightarrow\neg\neg\forall xP$$

taking $\exists y\forall z\,[T(x,x,y)\vee\neg T(x,x,z)]$ for P, we get a contradiction with Church's thesis. (However, it seems plausible that Σ_2 is closed under Church's *rule*.)

It would be interesting to see whether the function schemata used by Girard have models (for example the one consisting of the hereditarily computable terms) for which the interpretation of Church's thesis *is* satisfied. This would constitute a clear-cut advantage of the functions in [56] over the bar recursive functions.

Discussion. It cannot be stressed too much that limitations of the kind described do not spoil the *practical value* of interpretations. Quite generally, we must expect interpretations with a *familiar* range of functionals (which makes the interpretation easy to handle) to be *incomplete*, i.e., we must expect that $(\exists s\in F^-)A_0^-$ can be true even if A is not provable. After all, we know that the full complex of logical relationships is complicated: it would be foolish to expect a method (here: interpretation) which is both generally applicable *and* simple in cases of special interest to us.

Remark. To avoid misunderstanding, in connection with the topic of criteria for the *identity* of proofs, it is perhaps worth mentioning that we can easily introduce *some* equivalence relation between derivations even if we consider only *extensional* equality (between terms). Using the notation above, and an interpretation in $\mathcal{G}$, we define $\underset{\mathcal{G}}{\equiv}$ by:

$$d\underset{\mathcal{G}}{\equiv}d'\quad\text{if }s_d\text{ and }s_{d'}\text{ are }extensionally\text{ equal.}$$

For example, let us take the two derivations, say d and d', of

$$(p\wedge\neg p)\rightarrow(p\rightarrow p)$$

mentioned in the *Discussion* of §1a. Then, for practically any interpretation $\mathcal{G}$, we find $\neg d\underset{\mathcal{G}}{\equiv}d'$. Of course it would in general be absurd to suppose that

$$d\underset{\mathcal{G}}{\equiv}d'\Leftrightarrow\bar{d}=\bar{d}'\,.$$

We need a separate investigation, in terms of adequacy conditions, as in § 1(c).

(c) *Metamathematical principles* needed for understanding the syntactic transformations and functional interpretations. (More formally, one would speak of the principles needed to prove the assertions: Every derivation $d \in \mathcal{D}$ can be normalized, resp: If d derives A then some derivation d' of $\mathcal{G}$ derives $A_0(s_d)$.)

(i) In the case of *finite* derivations the principal point for syntactic transformations is to show that the *normalization figure* is well founded. As mentioned at the end of § 1a, this figure has derivations d at each node, and the immediate successors are the derivations obtained by applying a *single* normalization step to d, that is contracting any introduction of a logical particle followed in d immediately by its elimination. (The reason why, in general, we do not have a proof tree is that *several* such normalization steps can be applied to a single derivation.) Evidently if the derivations are finite, only finitely many steps can be applied, but a given derivation may be the result of applying a single normalization step to any one of infinitely many derivations, just as 0 is the result of a single computation step applied to $(s^n 0).0$ for $n = 0, 1, \dots$.

(ii) In the case of, possibly, *infinite* derivations there is the additional step of verifying that each (infinite) proof figure involved is well-founded. More explicitly, we give a *description* of a sequence of proof figures; as mentioned in (a) the functions used in this description are elementary. What has to be established, as indeed also in (i) above, is first that the figure described is *locally* correct, that is the formula at a node N of a proof figure is built up according to the rules from the formulae at the immediate successor of N; second in the infinite case, that the whole figure is well-founded; third that the normalization figure described is again locally correct (in the obvious sense) and, lastly, well-founded.

Evidently, *if* these normalization procedures are to be useful for consistency results, this will depend on finding *principles of proof of well-foundedness* which are, in the sense of § 1(e), more elementary than the principles that have led to the formal system studied; cf. also §3 below.

(iii) For the functional interpretations, at least where A_0 is not logically complicated, the principal point is *usually* the existence of functionals satisfying the axioms $\mathcal{G}$; "usually" in the sense that the proof of

$$(d \vdash A) \to d' \underset{\mathcal{G}}{\vdash} A_0(s_d)$$

is quite elementary for the current interpretations: note that "s_d" is the *name* of a term in $\mathcal{G}$; we do not use an enumeration functional for the functionals named in $\mathcal{G}$.

At the present time it seems fair to say that for the systems $\mathcal{G}$ used, such as Gödel's system T in [7], the *existence of the functionals involved is indeed evident*, but *not evident on particularly elementary grounds*. For hereditarily recursive objects (cf. HRO and HEO in Troelstra [46]), their natural definition uses the principles of first order intuitionistic arithmetic; and while the non-extensional operations HRO have more striking properties than HEO, properties which are easy to establish, the *principles* used are no more elementary. If $\mathcal{G}$ is interpreted by more specific computation rules (cf. App. I), the computability predicate has again a logically complicated form.

(d) *Digression* concerning the early work by Herbrand and Gentzen, on interpretations and syntactic transformations. Since I was probably the first person to "resurrect" Herbrand's work 25 years ago, to use its and generalize it [14], I can probably make a reasonable guess at its true and its superficial attractions. (Its limitations are evident to anybody who is not totally ignorant of the main stream of proof theory over the last 20 years; for example, at no place in SPT would it have been profitable to use Herbrand's formulation.) First and foremost, Herbrand's work involves refinements of model theoretic constructions; and since, when applicable, model theoretic reasoning is more easily visualized, Herbrand allows us the best of both worlds: model theoretic clarity and quantitative estimates (for a constructive explicit treatment). No analysis of the logical operations, in particular of the propositional operation of implication, is attempted. As described at the beginning of (b) above, Herbrand allows us to make clear the *content* of a theorem proved by restricted means without going into a detailed analysis of proofs. The true attraction is that often such information on the "constructive content" is all we want, and it is good to have a direct way of getting it. But, as so often, *il faut reculer pour mieux sauter*.

The first place where the striking superiority of Gentzen's analysis is evident is in connection with Heyting's predicate logic. Even if some kind of analogue to Herbrand's theorem can be squeezed out, cf. Minc [28], it is not elegant. Indeed, the basic feature of Herbrand's analysis, the separation of *all* quantifier inferences, is not altogether plausible here[5], since, in contrast to

---

[5] For footnote, see next page.

classical logic, implication involves much the same kind of abstraction as universal quantification; cf. §4 below. Thus once the intrinsic interest of intuitionistic systems for proof theory is recognized (as it has been by research over the last 40 years), we have here a definite limitation of Herbrand's analysis.

There is also a quite different extension of Herbrand's theorem provided by the *no-counterexample-interpretation* of [14]. Formally it is similar, not to Herbrand's own analysis, but to the proof given in Hilbert-Bernays by means of the ∈-theorems. Conceptually it is quite different. (The difference is that it uses functionals of lowest type and function *variables* in an essential way, concepts which Herbrand regarded as too abstract[6].) This is quite suitable for arithmetic and ramified systems, but not beyond, in the precise sense explained on top of p. 380 in SPT. And of course nothing corresponding to the normalization or even normal form theorems for full classical analysis have grown out of Herbrand's work.

3. Ordinal structures and formal theories of ordinals

In SPT, pp. 335-339 and 352b and, particularly, in [20] the ordinal structures used in proof theory, that is the orderings of the natural numbers and functions on them, were analyzed from an epistemological point of view. Since the discussions rambled on, let me summarize the main conclusion here.

As described in §2c, the main aim is to give an elementary justification of transfinite induction on (the domains of) our ordinal structures; in particular, more elementary or constructive than induction on abstract well-orderings. The key difference is that our orderings are *built up by* means of specific constructions (such as *addition* or *taking ω copies*); specifically by *iteration of these constructions only along previously introduced ordinal structures.* Consequently, transfinite induction on our ordinal structures can be analyzed explicitly in terms of the specific construction or build-up of our orderings; corresponding to the justification of ordinary, that is ω-,induction in terms of

---

[5] *Correction.* I did not recognize this state of affairs when I formulated the problem solved by Minc loc. cit. Note that, as usual, contraction of a disjunction $A \vee A \vee B$ to $A \vee B$ is not counted as an "inference". (The whole business of separating different *kinds* of inference by use of Herbrand-style theorems is still obscure; cf. SPT, 379 (iii) and footnote 39.)

[6] *Correction* (of Martin-Löf's [25], p. 12, 1.15). If one disregards this difference the complications of Herbrand's own formulation become quite incomprehensible (which is quite a separate matter from the fact that they can be avoided).

the construction of natural numbers by means of the successor operation. In contrast, in the abstract case, we have the operations *hidden* in the logically complicated assumption of well-foundedness. More formally, the abstract principle of transfinite induction is expressed by:

derive Ax from $[(\forall y < x)Ay \to Ax]$

which contains the logically compound subformula $(\forall y < x)A(y)$, and is therefore unsuited for a *logic-free metatheory*, required by the kind of proof theory here considered, cf. §2a(iii). The moment we have a *complete* set of build-up functions (in the terminology of Feferman [4]) denoted by, say, $f_1, ..., f_k$ we have a *quantifier-free* formulation of transfinite induction corresponding to the build-up of the ordinal structure; cf. [17], p. 172.

Let c_1 be (demonstrably) the *limit* of the segment generated by finite iteration of the f. Then we derive

$$x < c_1 \to Ax \quad \text{from } A0, \quad Ax \to Af_i x \quad \text{for } 1 \leqslant i \leqslant n.$$

To infer Ac_1, we need a replacement for the usual formulation $(\forall y < x)\,Ay \to Ax$ (of the progressive character of A) since it contains quantifiers. An evidently sufficient condition is provided by a derivation of

(*) $(\tau x < x \to A\tau x) \to Ax$

for some term τ already introduced. More generally the inference may be justified from the meaning of A, in particular (cf. [17], p. 172, 3.4214) if A is the property: the segment x has been built up by use of the operations $f_1, ..., f_n$; and conclude that the segment c_1 has been so built up. (More precisely, in the terminology of [20], c_1 has been *seen* to be so built up by methods implicit in the operations $f_1, ..., f_n$.) Thus, if the operations: $x \to f_i x$, preserve well-foundedness (in a suitable sense, in particular justifying transfinite iteration of a process along a well-founded ordering) we now permit iteration of the f_i themselves up to c_1 and, for $y < c_1$, denote the yth iterate by f_i^y.

Generally let c_{n+1} be the limit of the segment generated by iterating the f transfinitely often, but $< c_n$ times. Then we derive

$$x < c_{n+1} \to Ax \quad \text{from } (y < c_n \wedge Ax) \to Af_i^y x \quad \text{for } 1 \leqslant i \leqslant n, \text{ and } Ac_{n+1}.$$

These rules, in contrast to the abstract principle of transfinite induction above, are quantifier-free and hence logic-free if A is decidable. Note that the formulation above provides a more *explicit* analysis of the process of trans-

finite induction on $<$ than when (*) alone is regarded as sufficient to infer Ax. Here (*) is used only to infer Ac_n from $x < c_n \rightarrow Ax$.

In general, we need additional functions besides the build-up functions f, to establish the premises of the rules, in particular to express and establish the defining properties of the c_n. Thus the property *Lim*, of being a limit number, is defined by: $spx \neq x$ where s and p are the successor and predecessor function respectively; but to express in a quantifier-free way *that* this definition expresses the intended meaning we also need an additional binary g and a proof of $(spx \neq x \wedge y < x) \rightarrow y < g(y,x) < x$.

A particularly useful group of additional functions are the *inverses* of the build-up functions, also called *retracing functions*, as the predecessor is the inverse of the successor function. By use of such functions we can express explicitly the *fundamental relation* between

an element and its name built up from $0, f_1, ..., f_k$

that is, between an object x in the ordering and the term (built up from $0, f_1, ..., f_k$) which denotes x. This is fundamental because the term *reflects the construction of the segment up to x* as the numerals $(0, s0, ss0, ...)$, the *standard notations* for integers, reflect the construction of the objects they denote. An obvious consequence is this:

Let O_1 and O_2 be ordinal structures with isomorphic domains of ordinal α; both containing build-up and retracing functions as described above. Then O_1 and O_2 are isomorphic by the mapping *explicitly defined* as follows:

For x_1 in O_1, find its standard notation t by means of the retracing functions in O_1; by means of the build-up functions in O_2, determine the "value", i.e., the denotation x_2 of t in O_2.

In other words, our ordinal structures are determined *uniquely up to the explicit isomorphism above*. The reader may compare this *functorial analysis* of ordinal structures with the analysis of canonical representations discussed in subsection (d) of §1, an analysis which determines representations uniquely up to demonstrable equivalence.

Remark. People in proof theory, evidently interested in the aims descirbed in §2c, usually speak of the *ordinal* of a formal system; for example, ϵ_0 of first order arithmetic. Taken literally this conflicts with our discussion which shows beyond a shadow of doubt that, for their aims, not the ordinal nor even the ordering, but the *ordinal structure* is essential. But, practically speaking, no mistakes are likely to be made, just because the relevant ordinal structures are not only determined uniquely in the sense described, but isomorphic to the *familiar* ϵ_0-ordering (which is the ordering that people use without further analysis). The very *definition* of the ordinal ϵ_0 as the limit of $\omega, \omega^\omega, ...$ refers

to ordinal *functions*, essentially those needed for the ordinal structure brought out by closer analysis. Here it should be remembered that, at the present time, *we do not have a* **general** *scheme for associating in an intrinsic way an ordinal structure to a formal system* (but we do have one for associating an ordinal: bounds on provable Σ_1^1 - well orderings, induction principles needed for reflection principles; cf. SPT, pp. 340–341). Also, though such an intrinsic characterization would be aesthetically very attractive it is not needed for the specific metamathematical purposes of §2c which require only *some* ordinal structure built up in an elementary way.

In earlier publications ([20] but also SPT §12b) I considered the problem of associating ordinal structures or "natural" well-orderings to given ordinals. I shall not prusue this problem here, but two much more modest formal questions involved in setting up a *logic-free metatheory*, namely (α) and (β) below which were neglected in SPT.

(α) What are minimum *formal* requirements on quantifier-free systems of ordinal functions if they are to serve for formalizing the metamathematical principles discussed in §2(c)?

This question has of course perfectly good meaning if, as is usual in the literature, we have in mind a formal language for arithmetic and consider specific well-orderings of the natural numbers with number theoretic functions representing our ordinal operations. But since we are not concerned with their arithmetic properties and since we have just seen that the ordinal structures to be considered are unique up to a particularly elementary (explicit) kind of isomorphism we may just as well consider *formal theories of ordinals*. Without excluding the possibility that the variables may range over the particular (finite and infinite) sets which von Neumann used in his set theoretic ordinal structures, our theories will also be realized by ordinal structures that are defined by quite elementary (constructive) processes.

(β) Are there elegant formal theories of ordinals corresponding to the systems of arithmetic and analysis which are the principal subject matter of current proof theory?[7]

Remark. Since proof theory has been dominated so far by Hilbert's programme and other constructive aims, most of the few formal theories of ordinals

---

[7] In (c) we shall consider reasons why questions (α) and (β) were neglected for a long time.

studied in proof theory are constructive; in particular, they have models over "small" segments of the recursive ordinals, the function symbols being realized by (primitive) recursive functions on these ordinals, or, more precisely, their notations. But inasmuch as the requirement of constructivity is not always appropriate, as has been urged throughout this article, also other formal theories will be considered here, particularly in (b) and (c).

Naturally the latter are not directly useful for refining the kind of wellfoundedness proofs mentioned in §2 (c); specifically, refining them by eliminating the general notion of wellfoundedness in favour of operations on ordinal structures built-up in a particularly elementary manner. But apart from such an epistemological purpose, an ordinal analysis might be expected to have also technical uses, simply because it replaced the *qualitative* notion of well-foundedness by a *quantitative* formulation. To some extent this quantitative version, like interpretations in §2 (b) could give an answer to the recurring question:

  What more do we know when we know that a theorem can be proved by
  limited means than if we merely know that it is true?

  (a) We consider first question (α) applied to the theory of *syntactic transformations* described in §2(a).

  (i) If we are principally concerned with (infinite) proof figures, but not the conversion relation, a minimum requirement is this:

  For the natural coding of formulae the proof figures must be *definable* in
  the theory of ordinals and their basic properties formally derivable. The
  basic properties are of course that the proof figures are "locally correct",
  that is built-up according to the rules of proof under study, and that they
  are well-founded.

  Though the familiar normal form theorems for, say ramified analysis, by use of infinite normal (or cut-free) proofs have not been formalized in any theory of ordinals the method is fairly clear; for instance, we know already that the proof figures themselves have the appropriate order type (SPT, p. 332b). Indeed it was perhaps wise to wait because, as mentioned in (c) below, it is probably more elegant to deal with subsystems in the language of set theory rather than analysis: it would have been frustrating to labour over a formalization of the metamathematics of analysis.

  Let me here say a word about technical uses of a quantitative ordinal analysis of normal proof figures though *mutatis mutandis* related considerations also apply to an analysis of the conversion figures considered in (ii) below.

  For a sufficiently strong sense of "normal" it turns out that a bound on

the *ordinal size* of a normal deduction imposes a significant restriction on the theorem proved; for example ([41], p. 222, concerning normal derivations with the subformula property) a derivation of the well-foundedness of a relation has an ordinal exceeding that of the relation itself. (In the case of ordinary predicate calculus of first order, the necessarily *finite* length of a normal derivation is significant.) Here the order type of the normal form of any given deduction is a useful characteristic even if we do not go into the details of its ordinal structure.

For a genuine practical application, the considerations of §2(b) about "incomplete" interpretations apply. Quite evidently, the ordinal bound itself cannot possibly yield every independence result: as is well known, the addition of any true Σ_1^1 statement to a formal system does not alter its ordinal bound! (though it will in general alter its ordinal structure). But *if* an independence result is so brutal that it can be proved by looking at the ordinal bound only, the independence proof will generally be very easy.

(ii) The problems of formalizing the conversion procedures in §2a(ii) are less well understood; not surprisingly since, as mentioned already, those procedures have been neglected by proof theorists. For example, a satisfactory treatment (at least: satisfactory in the sense of SPT, p. 332(b)[8]) would require, above all, the study of the *computation* or *conversion figures qua* ordered structures. It is still not quite clear whether the *canonical* assignment of ordinals to these figures, say for first order arithmetic, really uses all ordinals $< \epsilon_0$. Of course, we know from §2(c) (i) that a *proof* of well-foundedness in a theory of ordinals (with successor, addition, exponentiation) needs induction $\leqslant \epsilon_0$. But this is a separate question from determining the relational type of the figures themselves or the least ordinal into which they can be embedded; more precisely, embedded by means of mappings which are characterized by *definability* (not also *provability*) conditions.

Evidently the problem of *formalizing normalization theorems in a theory of ordinals* induces minimum requirements on such a theory of ordinals in a way which is exactly analogous to (i) above.

It is perhaps not superfluous to go into the *nature* of "the" problem of formalizing say Gentzen's original consistency proof. This will explain and, I think, even justify people's reluctance so far to work on this problem. - As

---

[8] Recent experience seems to confirm this dogma: in [9], Howard considers a formally similar situation, namely a computation figure and assigns ordinals (such that the application of a reduction step reduces the ordinal). The assignment is not, explicitly, related to *the order type of anything*, and is certainly not pleasant to check.

somebody said: A formalization is the outward and visible sign of an inward
and invisible grace, namely of one's understanding, of knowing what one is
talking about. (Formalization shares both the advantages and disadvantages, for
a given person, of correct formal manners.) Specifically, a *principal* difficulty
is to decide in what terms, in what *language*, the formalization is to be given, a
matter which is certainly not decided by the informal argument; and then, to
decide to which points special attention should be given. Now, it is fair to say
that though it was immediately clear that some striking reduction had been
achieved by Gentzen, not even the most fundamental and now very familiar
point was clearly realized: that ϵ_0-induction is applied to a quantifier-free,
that is logic-free predicate. If one does not have reasons to *look* for a quanti-
fier-free formalization, one will truly end up with the splendid joke: Gentzen
proved the consistency of ordinary, that is ω-induction by means of ϵ_0-induc-
tion. Equally to the point is an example where somebody stumbles on a
formalization, but does not profit from it. As already stressed in SPT, top of
p. 332, Schütte [41] formulated explicitly the properties of some ordinal
functions needed for the consistency proof. But the significance of these
properties did not spring to the eye from the formalization; it was found only
when one *looked* for it, in terms of an intrinsic characterization of the natural
ordinal structure on ϵ_0.

   (b) By §2c (iii), the relation between ordinals or ordinal structures and
the functions (of higher type) used in the interpretations of §2b must *in gen-
eral* be expected to be tenuous. More precisely, the theory of ordinal struc-
tures is of course directly relevant if the functions are thought of as rules for
applying given computation or reduction procedures; in this case the minimum
requirement on such a theory is completely parallel to that of a (ii); the com-
putation figure should be definable (or: computation tree if the rules are deter-
ministic) and its basic properties should be provable. We shall take up this
matter in App. I.
   But, as a matter of historical fact, the functions one had in mind when in-
troducing the no-counter example-interpretation or Gödel's [7], were not
thought of as such computation rules. In other words the existence of the
functions involved was evident on other grounds. Certainly, since then, rela-
tions between those functions (or at least schemata satisfied by them) and
ordinal structures have been established; cf. SPT, pp. 332-333 for the con-
tinuous functions used in the no-counter example — interpretation and
Howard's assignment mentioned in footnote 8, for Gödel's T. However, by no
stretch of the imagination could it be claimed that these relations provided
the original reason for introducing the interpretations; indeed Gödel explicitly

proposed the principles formulated in T as an *alternative* to Gentzen's use of ϵ_0-induction, and hence independent of the latter (but possibly less elementary, cf. §2ciii).

The functions of higher type referred to in the last paragraph have the natural numbers as ground type. More recently (cf. (c) below), functions of higher type over the ordinals have been studied; naturally with a very direct relation between functionals and ordinals.

(i) For the case of continuous (and hence of course extensional) functionals of *lowest* type over the integers, with the discrete topology on the integers and the product topology on functions, we do have a close relation, since there is an *inductive generation* of this class of *functionals*[9]. But, as just mentioned, many of the schemata in the literature for such continuous functionals are introduced for other reasons, in particular because their computability is made evident by use of the notion of choice sequence. In such cases, the computation "procedure" we have in mind *does not refer to the inductive generation but to a method of trial and error*. With the usual convention (e.g. [16]) we take a neighbourhood function f_0 on finite sequences of natural numbers, a choice sequence α and try out $f_0(\bar{\alpha}0), f_0(\bar{\alpha}1), f_0(\bar{\alpha}2)$ till we hit an m such that $f_0(\bar{\alpha}m) \neq 0$; cf. computations on those definitions of recursive functions which are given in Kleene's form $U[\mu_y T(e_0,n,y)]$, where, for each n, we try out $T(e_0,n,0), T(e_0,n,1), ...$ till we hit a $T(e_0,n,m)$ which is true. For *these* definitions (not to be confused with the *computational* normal forms in [42] or [44]!) the obvious order type of the computations has a finite ordinal; cf. also App. la(i).

The simple general point made in the last paragraph is, perhaps, further supported by detailed work. On the one hand we have the easy proof of computability in [42] pp. 225–227, of the closed terms, actually for higher types, of Gödel's theory T [7], when they are thought of as denoting computation procedures formulated in the theory itself, discussed further in App. I (a). On the other we have the complications (in [44]) arising from bar recursion of type 0, a schema which is suggested by the notion of choice sequence. At the present stage we do not know if the difference is intrinsic, see also the end of Troelstra's paper [46].

---

[9] In [23] and elsewhere it was overlooked that Kalmar discussed this fact some 15 years ago in [11]. (In the logical literature this relation is more familiar from the order type of the unsecured sequences of such functionals in the socalled Brouwer-Kleene ordering.)

(ii) For familiar functions of higher type (or, more precisely, functions satisfying *topological* conditions such as the socalled continuous ones of [16] or [40]) we do not know any simple connection with ordinals, in particular, no inductive generation similar to (i).

To be precise, by Gödel's theory of constructible sets we have a completely systematic method of connecting *any* theory of functionals which can be developed from the usual axioms of set theory (or additional axioms which hold in L) with a theory of ordinals. We need only restrict ourselves to the constructible functionals and formulate the latter as, e.g., in [48]. But this goes well beyond the principles envisaged in existing proof theory (on the analysis of SPT, p. 351, §12).

(iii) Quite recently Feferman [6] has made very interesting use of suitable *segments* of the constructible hierarchy to establish connections between functionals and ordinals (or more precisely number theoretic functions defined by means of functionals of finite type). The crucial differences are, firstly, that the functionals considered are *not constructive*, not even in the weak sense of having a recursive realization for the *function* symbols (i.e., not even constructive in the sense of the theory of species); secondly, ordinals do not enter in the form of the ordinal structures here described but simply as the *ordinals of hierarchies*.

Since the formal theories of the *particular* hierarchies which he considers, up to ϵ_0 steps, have a proof theoretic analysis by quite elementary means, his work can serve as an *intermediate* stage for constructive proof theory. (From the point of view of method this is to be compared to intermediate uses of model theory described in SPT, Note II and, particularly, his own work in [51]; the difference is that not models of quantification theory, but suitable classes of extensional functionals are introduced.) Nevertheless, in the sense in which I meant the *warning* on p. 334, 1.-14 to 1.-12 of SPT, it is now necessary to warn against this warning.

(c) *Theories of sets and ordinals.* Traditional proof theory studied mainly formal systems formulated in the language of arithmetic or analysis (with variables for natural numbers and sets of them). This choice was natural when one wanted to use constructive methods and hence to avoid abstract assumptions. Since the most familiar uses of sets and transfinite ordinals were highly abstract, the mere use of the corresponding languages was thought to carry a risk of confusion. Times have changed; we have a better grasp of principles and we are familiar with quite elementary models of axioms in the languages of sets

and ordinals. Consequently it makes good sense also in proof theory to use these elegant languages which, in particular, avoid elaborate coding procedures.

(i) Relations between subsystems of the usual set theories and (the usual) systems of analysis were discussed in (SPT), Technical Note VI, pp. 375–376. Some set theories corresponding to subsystems of analysis were considered in (SPT), p. 376(c), but more interesting ones have been studied since then. For example, Feferman's system [5] is intended to provide a set theoretic system corresponding to his system IR of analysis [3] or, equivalently, to his hierarchy of ramified analysis; or [10] where an elegant set theoretic version of "bar-induction" is treated without any (stated) systematic purpose.

(ii) Relations between set theories and theories of ordinals: A non-recursive step. The obvious first step, already mentioned in b(ii), is simply to specialize the theory of constructible sets as formulated in [48] to subsystems of usual set theory. By standard collapsing arguments, we get minimum well-founded models.
   In general these minimum models are non-recursive in the following sense. Either the *domain* itself contains non-recursive ordinals, or, if it is a segment α of the recursive ordinals, the *ordinal functions* are not realized by recursive functions on the natural well ordering of α. (Equivalently, the ordering relation is not recursive on the domain made up of the *terms* of the theory.)
   As an example of the latter Zucker studied the minimum model of group 1 in [45] (in the course of analyzing the claims of Wette [47] regarding this group of axioms). The domain consists of the ordinals $< \omega^\omega$, but it is easily seen that the functions named in the axioms are *not* realized by recursive functions on (the canonical ordering of) ω^ω; cf. the review of [47]. As an example of the former, Prawitz [35] has given an elegant theory of ordinals (and specific functions on them) which corresponds to full analysis; here the *domain* itself of the ordinals needed goes beyond the recursive ones.

(iii) Relations between set theories and theories of ordinals: recursive models. More explicitly, we now look for theories of ordinals which are realized by segments of the *recursive* ordinals $< |a|$ (where $a \in O$ [12]) such that the ordinal functions on the ordering $\{(x,y):x<_O y<_O a\}$ are also recursive. Most of the work along these lines has been done since (SPT) was written though Howard introduced a first example some years ago in Section VI of the privately circulated Stanford Report on the Foundations of Analysis. The methods so far used proceed somewhat indirectly as follows. A quantification theory of ordinals and particular species of ordinals is set up which corresponds

to the subsystem of set theory or analysis considered. For example, for the
ories codifying the familiar principles of *generalized inductive definition* (g.i.d.)
we have an absolutely standard reduction to the theory of ordinals: to find the
least set X_A satisfying $\forall x [A(P,x) \to P(x)]$, for monotone A, we put $P_0(x)$
$\leftrightarrow A(\perp,x)$, and define by transfinite recursion a sequence P_α for ordinals $\alpha > 0$
by

$$P_\alpha(x) \leftrightarrow (\exists \beta < \alpha) A(P_\beta, x)$$

where $x \in X_A \leftrightarrow \exists \alpha P_\alpha(x)$.

This quantification theory is *interpreted* in a suitable system of higher type,
by Howard *loc. cit.* for non-iterated g.i.d., in the style of Gödel [7], by Zucker
[49], for iterated g.i.d., using the modification of 3.5.1. in [16]. "Suitable"
means here that the closed terms can be used as the domain of a manageable
model. For readers familiar with Gödel's proof that the continuum hypothesis
holds for L (the collection of constructible sets), the current methods, first
made explicit by Feferman [4], are best described as follows. As in b(ii) we
think of the theory of constructible sets formulated in a theory of ordinals
[48], with an ordinal theoretic relation $\alpha e \beta$ to mean: the set with number α
is a member of the set with number β.

The first step is to set up a quantifier-free system which is evidently satis-
fied by a "large" (uncountable) segment of the ordinals when the function
symbols including constants are realized by familiar ordinal functions: call
this structure 0^∞. This corresponds to the step in Gödel's "collapsing" argu-
ment where some large ordinal is adjoined to a given segment $< \alpha$ (where α is
a cardinal in L), and the "Skolem hull" is formed. The latter does *not* fill up a
segment of the ordinals. We may regard the formation of the Skolem hull as
the construction of a *term model*, consisting of constants for each ordinal $< \alpha$
and for the adjoined ordinal, and terms formed by means of function symbols
f for the Skolem functions used. By definition, the realization $\bar{f}$ of f in a term
model (also called: canonical realization) is defined by the action

$$\bar{f}: t_1, ..., t_n \to ft_1...t_n$$

for all *n*-tuples of terms (if f has n arguments).

The next step in the collapsing argument is to consider the ordering of the
terms

$$\{(t_1, t_2) : t_1^s < t_2^s\}$$

where t^s is the realization of t in the Skolem hull, that is the ordinal denoted by t. The two facts essential for Gödel's proof are that the ordinal α_1 of this ordering has cardinal α (in L) and that the ordering is definable in L_α. This ensures that the Skolem hull *together* with its structure is definable since there is an obvious indexing (in L_α) of the terms and hence of the canonical realization of the function symbols; in particular, $\{t_1, t_2 : t_1^s < t_2^s\}$ is also definable. (Evidently, the Skolem hull regarded as a set theoretic object and not only as a structure is not in general definable in L_{α_1}.)

For the present prupose of finding recursive models, analogous but more delicate requirements are needed. Instead of preservation of cardinals we want α_1 to be recursive (for the given α). And instead of (invariant) definability over L_{α_1} we want the relation:

(*) $\qquad \{(t_1, t_2) : t_1^\infty < t_2^\infty\}$

to be recursive where now t^∞ is the ordinal denoted by t in 0^∞. Evidently, if this is satisfied, once again the function symbols are automatically realized by recursive functions (on these terms).

Feferman [4] gives some useful general conditions on ordinal functions for which (*) is satisfied. Actually in [4] his primary concern is with the more *elementary* parts of proof theory where no collapsing occurs at all (for an application of collapsing see, e.g., Zucker's thesis [49]). We start again with a quantifier-free system which is evidently satisfied by familiar ordinal functions over familiar structures. But now it turns out that the ordinals denoted (in these familiar structures) by the terms fill up a segment anyway, i.e., the systems are *replete* in the notation of [4], and (*) holds too. In short, we have an *absoluteness* or *invariance* property:

The terms have the same value whether the function constants get their intended (familiar) or their canonical realization.

An interesting difference between iterated and non-iterated g.i.d. is just this: in the latter case, the closed terms used in the interpretation [49] do not have this invariance property.

Remarks. Howard's interpretation is quantifier-free and the model is recursive in the strong sense that the = relation (even) between terms of non zero type can be realized by a recursive relation. Zucker's interpretation allows indeed a recursive model for the domain, the functions and even the = relation, but has a *residue of logic* in the form of two species which are not realized by recursive predicates. (This asymmetry between functions and predicates is of course quite familiar.)

From the point of view described at the beginning of this section, of the *build-up* of ordinal structures, the use of functionals of higher type over the ordinals is quite natural. A functional of higher type corresponds to reflection *on* the process of iterating operations of lower type (though it does not make explicit the details of this process). Some *formal* notions and axioms express-ing this use of higher types can be found in one of Feferman's articles in the Buffalo volume [52]. A more delicate analysis, which confines itself to oper-ations on ordinal structures (of lowest type), was given by Howard [53], and Zucker [49] in terms of the particular build-up functions employed in Bach-mann's hierarchy.

Evidently, depending on one's interest, further work will pursue two, at least *prima facie* diametrically opposite aims. On the one hand the work will be *extended* using, no doubt, principles suggested by familiar set theoretic material such as Mahlo's iteration procedures. So to speak the processes are the same, only the basic operations (taking successors in the constructive case and power sets in the set theoretic case!) are different. Indeed from this point of view, ordinal constructions that are here developed should serve as a *testing ground* for problematic axioms of infinity. In fact it may even turn out that the term "axiom of infinity" is not quite appropriate, that is, that (cardinal) *size* derives essentially from the basic operations, and that the full *complexity* is naturally present in the recursive case; cf. footnote 37 on p. 377 of SPT[10].

On the other hand one will examine the models constructed more carefully and try to find concepts in terms of which essential differences can be analyzed precisely. For example I am not yet persuaded that the difference between iterated and uniterated g.i.d. is negligible even assuming, as seems reasonable, that on interpretation of the logical operations is found which is really appro-priate to the use of g.i.d. (as explained e.g. in SPT, § 12(a) or p. 352). If this is so, closer inspection of the functionals used in the interpretations by Howard and Zucker may help one find the concepts needed to state the difference.

4. Logical operations: proofs and functions

Nearly 40 years ago Heyting gave a meaning to the basic logical notions (of proposition, species and operations on them) [8]. The meaning was explained, as I read it now, in terms of *two* primitives, namely *judgement* or proof, speci-fically proofs of identities which are of course "logic free", and *functions*

---

[10] The qualification "naturally" is essential since size is always eliminable in the sense that there are denumerable elementary submodels.

whose arguments and values are hereditarily pairs (p,π) where p and π denote proofs and functions resp. For example, (p_0,π_0) establishes $A \to B$, for short

$$(p_0,\pi_0) \vdash A \to B$$

means that p_0 is a proof of

$$[(p,\pi) \vdash A] \Rightarrow (\pi_0(p,\pi) \vdash B]$$

for all p and π. Heyting also wrote down formal laws which can be seen to be valid (even) from a very rough understanding of the two primitive notions of proof and function.

Now the single most important point to remember is this.

Though Heyting *intended* p to range over constructive proofs and π over constructive functions, what he *actually* said about these primitives, the laws he actually asserted, made little use of any detailed analysis of the qualification *constructive*. Heyting certainly emphasized the fact that *any* list of formal laws leaves some room for different interpretations; but he failed to stress just how much room was left open by his particular list of formal laws.

(a) As is to be expected, for a more detailed analysis of constructivity, specifically if constructive number theoretic functions are recursive, Heyting's laws are incomplete, and the set of valid laws (for the language of predicate logic) is not recursively enumerable [19][11].

Though it is very natural to present various kinds of "realizability interpretations" (discussed at this meeting by Troelstra [46]) as variants of Heyting's scheme (cf. [17], p. 128–129, 2.31–2.32) this is not usually done. Also interpretations of the logical operations "in terms of" formal derivability such as Prawitz' [34], fit in here where p now ranges over (proofs described by) derivations in a given system, and π over a (very restricted) class of functions. For yet another variant (used in the first of the recent discussions of Heyting's scheme [15]), the "guiding" model let p range over proofs which are, hereditarily, formalizable in a progression of formal systems, and π over functions such that, in the definition of $\to$ above, the level of (the first element of) $\pi_0(p,\pi)$ was not too far from the level of p; cf. App. Ic(ii).

---

[11] *Correction.* Contrary to Myhill's facile explanations, [29], p. 327, l.14, incompleteness is not due to lack of analysis of the meaning, but to knowing what we are talking about; cf. the case of higher order logic where we have incompleteness (of formal systems) if we do specify what we are talking about, namely principal models.

Remark. It is fairly easy to generate paradoxes if one tries to set up a formal theory of constructions (i.e., of proofs and functions). But, as usual, this indicates equally that we have too *many* reasonable interpretations as that we have none! Indeed, if we have anything in mind at all when setting up formal laws which are inconsistent, this means that one law is valid or at least plausible for one meaning, another for a different meaning. If we had no meaning in mind the proposed laws would not be plausible at all.

(b) To give a positive turn to the "inadequacies" noted in (a), it is of course necessary to find intrinsically interesting or at least useful examples among the unintended interpretations. The examples mentioned in (a) involve *restrictions* on the notions intended by Heyting since (cf. SPT §13) proof theory has been traditionally interested in particularly *elementary* methods; its aim was to be "reductive" (in the sense of Prawitz [36])[12]. But good use can be made also of non-constructive interpretations. For example, Feferman's work on "predicativity relative to the notion of natural number" or, more simply, on "reducibility to arithmetic" uses the non-constructive meaning of numerical quantification, and thus the functional E [13], for function variables α

$$E\alpha = 0 \leftrightarrow \exists x(\alpha x = 0) , \qquad E\alpha = 1 \leftrightarrow \neg\exists x(\alpha x = 0)$$

is quite appropriate. This work in [6], already mentioned in §3(b) is also formally very satisfactory.

(c) *Explicit definitions.* There is one use of Heyting's formal laws which does not come under his scheme at all but is, I believe, much closer to the interests of mathematical practice. The general consideration is this.

In ordinary classical mathematics the operations $\exists$ and $\vee$ are so to speak unemployed. They are explicitly definable in terms of $\neg$ and $\forall$, resp. $\wedge$. So we ask for formal laws that reproduce ordinary classical reasoning for the so-called negative fragment $\neg, \wedge, \forall$, but for existential theorems and disjunctions we have a stronger requirement than truth, i.e., we require (for closed formulae)

$\exists x A$ is derivable (if and) only if, for some term t, $A [x/t]$ is derivable.

$A \vee B$ is derivable (if and) only if either A or B is derivable[13]

---

[12] It goes without saying that we all know one interpretation which goes *beyond* Heyting's, where provability is identified with truth, the pair (p,π) reduces to the single element π, and *all* functions are considered!

[13] In systems including e.g. arithmetic, the second condition is included in the first since $(A \vee B) \leftrightarrow \exists x [(x=0 \rightarrow A) \wedge (x \neq 0 \rightarrow B)]$.

NB. These conditions are *not* (even) required by constructive validity, e.g. of predicate logic, contrary to an almost universal misconception! Consider formulae A and B of predicate logic with, say, a single predicate symbol X; suppose X is binary.

$$A \vee B \quad \text{is valid if and only if}$$

for all species D ("D" for domain) and all $X^* \subset D \times D$, $A^* \vee B^*$ holds when the variables of A and B range over D and X is replaced by X^*; for short

$$\forall D(\forall X^* \subset D^2)(A^* \vee B^*) .$$

But it is certainly not evident that therefore

$$[\forall D(\forall X^* \subset D^2)A^*] \vee [\forall D(\forall X^* \subset D^2)B^*] .$$

The requirement for existential theorems involves an equally if not more startling uniformity

$$\forall D \forall X^* \exists x^* A^* \rightarrow \exists t \forall D \forall X^* A'$$

where A' is obtained by replacing x^* in A^* by the value t^* of the term t in (D, X^*). For first order formulae A, t does not contain (a symbol for) D nor X and t^* does not depend on D nor X. For higher order logic using abstraction terms, t^* involves D and X, but t does not contain a symbol for D.

The present "non-logical" use of Heyting's laws is particularly interesting in connection with various formulations of the *axiom of choice*. As has often been remarked there is an almost universal misconception about the relation between the axiom of choice and explicit definability, because of the accident that in the usual formulations of the axioms of set theory all axioms other than the axiom of choice assert the existence of sets for which we have an explicit definition. What is overlooked is that there are existential theorems without explicit realizations which are classical *logical* consequences of these axioms (SPT p. 349 bottom). In other words, *unrestricted use of* (ordinary) *classical logic does not preserve explicit realizability of existential theorems* (more precisely, *demonstrable* realizability; cf. the problem top of p. 372 of SPT).

In contrast, if we use Heyting's rules and even if we add the axiom of choice (and also the negative translation of the axiom of choice! cf. [39]) we have realizations of existential theorems.

Remark. One reason for supposing that these questions of existential realizations are closer to the interests of mathematical practice is simply that the central problem of interpreting →, which has been discussed very clearly over the last 40 years, has simply not caught "public" attention. Questions of existential realizations have. More theoretically, the latter questions are intelligible without logical preoccupations, the central question is not. It would be interesting to know how one would have set up *discovering* suitable formal laws which ensure realizations of existential theorems (without having in mind Heyting's own interpretation). Amusingly, Heyting's laws are naturally set out in such a way that they are *exactly* the classical laws except for the additional rule

$$\neg\neg A \rightarrow A$$

which does not mention existential quantification at all!

(d) *Other uses.* More problematic, but also more ambitious, applications of Heyting's formal systems rely on Kripke's discovery [34] that they are valid for an interpretation of the logical operations suggested by the idea of a *potential* totality or of an ever *expanding* body of knowledge. I say "suggested by" and not "in terms of" because the *sequence* of models used to represent the expansion of the universe (or of our knowledge) is itself usually treated in the literature as an ordinary set. This procedure provides no epistemological analysis at all. An alternative is to regard the logical operations applied to potential totalities as *primitive*: Kripke's interpretation is not regarded as an explanation or "reduction" but as asserting properties of these logical operations obtained by reflection upon their meaning. From such an impredicative view of these operations it is perfectly proper to use them in formulating their own "semantics".

Several people including Kripke (cf. the reference in Putnam [38, 284]) or Pozsgay [30] have proposed to use Heyting's rules for an analysis of the idea of generating sets by *arbitrary* iterations of the power set operation (cumulative hierarchy of types). Tacitly one assumes in ordinary set theory that one is dealing with a *segment* of the hierarchy since otherwise the classical interpretation of the universal quantifier does not apply here, there being no set of all sets; cf. [17] p. 101, 1. 19–22 and footnote 12 on p. 120. The situation is strictly analogous to the simple-minded finitist's idea of generating *all* hereditarily finite sets. Also formally the treatment is parallel provided only one uses heavily *set induction* corresponding to ordinary induction in arithmetic; for example, in arithmetic one proves the decidability of equality by induction

from the successor axioms and then goes on to prove the decidability of for-
mulae containing only bounded quantifiers, having introduced suitable arith-
metic functions; similarly, by means of set theoretic induction one *proves* the
decidability of formulae containing only bounded quantifiers from the deci-
dability of atomic formulae by essential use of the elementary set theoretic
operations, corresponding to the "suitable" arithmetic functions in arithmetic.

The *comprehension axiom* must be modified to

$$\forall a \forall x_1 ... \forall x_n [\forall x(A \vee \neg A) \rightarrow \exists y \forall x(x \in y \leftrightarrow [x \in a \wedge A])]$$

where the free variables in A are among $x, x_1, ..., x_n$ and y does not occur in
A. The restriction to decidable A is required because for the present meaning
of *set* (the cumulative hierarchy), a set is a well defined unity (in Cantor's
words) with a definite extension.

In [54], Note 1 §2(b), I suggested the *reflection principle* in the form

$$\forall x_1 ... \forall x_n [A \rightarrow \exists \alpha (x_1 \in V_\alpha \wedge ... \wedge x_n \in V_\alpha \wedge A_\alpha)]$$

for an arbitrary A with free variables $x_1, ..., x_n$, where α ranges over the or-
dinals, V_α denotes the segment of the hierarchy up to α and A_α is obtained
from A by restricting all quantifiers to V_α.

It should be noted that these modified schemata imply the corresponding
familiar axioms when the "negative" translation is applied (cf. SPT, p. 344
(iii)). Thus *the present interpretation justifies the use of classical logic applied
to the axioms listed.* Incidentally, I see no justification (for the present inter-
pretation of the logical operations) of the classically equivalent form of the
reflection principle above:

$$\forall x_1 ... \forall x_n \exists \alpha [x_1 \in V_\alpha \wedge ... x_n \in V_\alpha \wedge (A \leftrightarrow A_\alpha)]$$

that is, of finding α for given values of $x_1, ..., x_n$, merely from the form of A
without knowing whether A is well determined; only its negative translation is
justified. But not even this much is apparent for the "full" principle
$\exists \alpha (\forall x_1 \in V_\alpha)...(\forall x_n \in V_\alpha)(A \rightarrow A_\alpha)$; cf. footnote 6 on p. 10 of [57].

The reader of Note I §2 b in [54] will have noticed that an argument for
the reflection principle must use specific properties of the concept of *set*, not
merely the fact that the ranges of our variables are indefinite. These properties
cannot be shared by (the finitist's idea of) the indefinite hierarchy of heredi-
tarily finite sets, since the reflection principle does not hold for the latter.
Presumably the essential point is that the finitist can prove A by use of the

generation principle: $x, y \to x \cup \{y\}$ (in his language under which no finite V_α is closed. (Here we have another example of an assertion about sets which cannot be said to be "transferred" from our knowledge of finite sets; cf. [54] Note I §1 b, especially footnote 9.)

Remark. What is problematic here is not primarily the validity of Heyting's rules. As has been stressed repeatedly they tend to be valid whenever we speak, even quite vaguely, of processes. The problem is whether, when dealing with a *particular* process, Heyting's schema is fruitful; cf. footnotes 16 and 19 to Note I of [54] or footnote 3 of [20]. For example, Schütte [41], following Ackermann's ideas (cf. [1]), discusses not unreasonable logical laws for "indefinite" properties for which $A \leftrightarrow \neg\neg A$.

Appendix I
Computations and formalist semantics of the logical particles

The topics treated here are not yet ripe for a systematic exposition, though there has been renewed interest in them recently (see [44] or Prawitz' paper in this volume); following related earlier work by Curry and Lorenzen, in particular his operative Logik. At the present stage it seems useful to ask 3 questions which have been neglected in the literature.

(α) What are the *general* aims of this kind of work?

(β) What are the *exact* relations between this work and more familiar material in the literature?

(γ) For which (mathematical) results is this work actually *needed*?
Rules (or functions) are considered in (a) and (b) below, logic in (c).

Ad (α) it seems clear that the general aim is to replace our usual impredicative notions by notions satisfying the basic *formalist requirement*:

All objects involved should be explicitly listed.

This kind of reduction is similar to but more radical than that considered in [19] where the abstract impredicative notions are replaced by *recursive* ones. The difference is that though we have of course a 'listing', we do not have an *internal*, that is recursive, listing of recursive definitions. The steps from the abstract to the recursive versions are given in a(i) and c(i) for computation and logic resp.

Ad (β), the 'familiar' material in question consists of course of the recursive versions above. The general principle for using these, formalistically not acceptable versions is quite commonplace:

Formulate explicitly the closure properties of the class of recursive operations needed to verify the specific laws considered (which hold for the abstract interpretation).

In other words, explicit formalization is good for a formalist foundation. As a corollary: *some* formalistically acceptable interpretation is usually easy to come by from known work, and the principal fruitful problem for current research is to analyze what further requirements a formalist foundation should satisfy: not the existence of some (coherent) formalist interpretation, but the choice among such interpretations is problematic. This matter will be considered in b(iii) and c(iii) in terms of the *expressive power* (of simple languages) provided by formalist interpretations. Put differently: here we need not adopt the formalist *philosophy*, of mathematical precision and

mathematical reasoning generally, which requires formalist interpretations; but we expect the latter to be useful in areas which do have formalist character. For example, the preparation of any mathematical problem for a computer constitutes some kind of formalist reduction; it does not seem unreasonable to expect that there is a useful general theory here comparable to model theory which is a general theory of axiomatic systems and which has turned out to be useful in mathematical practice.

Ad (γ) we compare mainly results for the particular system *T* (introduced by Gödel [7]) obtained by use of the 'crude' recursive model HRO in [46], already mentioned in §2c(iii), and of the formalist interpretation in [44]; cf. b(ii). In fact the principal fruitful problem here is to *search* for (interesting) problems which do need the more refined formalist interpretation.

Besides these general points (α)– (γ) which have analogues in almost any subject at a corresponding stage of development, two *specific considerations* are introduced:

Firstly, the distinction between those formalist foundations which have an *abstract impredicative residue in the metamathematical explanations* and those which are, roughly speaking, hereditarily *logic-free*; see the introduction of [19] for the corresponding situation with recursive versions, especially the logical residue in the quantifier combination ∀∃ used to define 'recursive'. Here b(i) and particularly c(ii) are relevant.

Secondly, as already mentioned at the very beginning of this article, I propose to use formal relations, which were discovered in work on a formalist foundation, *for the ananlysis of the abstract notions themselves* cf. c(iii). This is not to be regarded as a kind of *tour de force* (prompted by my distrust of formalist philosophy above), but simply as an instance of traditional procedure. For example, when we study the abstract notion of *set*, we also consider the collection L_ω of hereditarily finite sets generated from ϕ by the operation: $x, y \to x \cup \{ y \}$, and formulate axioms and definitions valid for L_ω. But once our attention has been drawn to these axioms we make a fresh start and ask if they also hold for the abstract notion; cf. note 13 on p. 146 of [19].

I begin with computations (and not with logic) because here, at least occasionally, our *usual* meaning is formalistic; specifically, when we write down the formula

(*) $0 + s0 = s0$

(where *s* is the successor symbol) we sometimes *mean* that the formula (*) itself is obtained by *applying the defining equations for addition*

$$a + 0 = a \qquad a + sb = s(a + b);$$

here it is tacitly understood that the LHS in a defining equation is replaced by the RHS (and not vice versa) until a closed term is reduced to its *canonical form*, namely a *numeral* $0, s0, ss0, ...$ Thus the defining equations play a *double role,* as *computation rules* in the sense just described, and as *assertions* about the mapping, from numerals to numerals, determined by these computation rules. The distinction is genuine; for example, the totally undefined function is computed from the rule: replace fa by $2fsa$, but the *constant* function $\dot{0}$ satisfies

$$fn = 2f(n + 1) \quad \text{for all } n.$$

The rest of this Appendix will be quite concentrated, assuming a good knowledge of the literature cited.

(a) *Semantics of T* here, as in [7], regarded as an *equational calculus.* Thus, as is familiar from the reduction of primitive recursive arithmetic to equational form, we need an *equality functional* (here for each finite type σ) and formulate *induction* by the rule:
For terms t_1 and t_2 of type $0 \rightarrow \sigma$, derive $t_1 n = t_2 n$ if $t_1 0 = t_2 0$ and $t_1 sn = tnt_1 n$, $t_2 sn = tnt_2 n$ have been derived for some term t of type $0 \rightarrow (\sigma \rightarrow \sigma)$, with variable n of type 0.
    Then an interpretation or 'realization' must give a meaning to the constants, the variables and $=$ only; for a *constructive* treatment it would be natural to require also a specified class of *proofs* (of formulae with free variables) as in [22], p. 202, 1.-11. While at some later stage it will perhaps be useful to consider *all* models of T, for the present purpose of answering questions $(\alpha) - (\gamma)$ at the beginning of this Appendix, it is crucial to pick out *relevant models.*
    The reader is referred to [7] for the meaning which led (Gödel) to the formulation of T, in terms of the abstract impredicative notions of *constructive function of finite* type over the natural numbers and of *definitional equality.* Contrary to the most familiar idea of function of finite type in constructive mathematics, even for $\sigma \neq 0$, an operation of type $\sigma \rightarrow \tau$ **does not operate on 'approximations' to the argument (of type σ)** but on a rule (or *definition* of the argument). Thus we do not expect extensionality and, a fortiori, not continuity. Any *general doubts* about the sense of these explanations are quickly dispelled by its recursive version HRO, developed by Troelstra in [46], already mentioned in §2c(iii).

(i) Here the *domain* of the model of *T* consists of HRO, the *hereditarily recursive operations*, that is (physical) *programmes* of a Turing machine, which are, hereditarily, well-defined; the relation realizing = is *literal identity*. Closed terms are interpreted by a programme in such a way that the axioms of *T* hold in this model. For reference in (b) below: we can get non-mechanical models too, for example hereditarily hyperarithmetic operations.

Evidently, not every object in HRO has a *name* in *T* for this interpretation, which is unacceptable in a formalist foundation; HRO is not recursively listable. The following *refinement* is typical of the use of recursive models referred to in the introduction:

Since *T* is quantifier-free we may cut down HRO to those programmes which *do* have names in *T* (and inspection shows that we can find a primitive recursive list for each finite type). Since each programme in HRO is well-defined for *all* of HRO it is automatically well-defined for our subclasses.

Discussion. While the results of [46] to which we return in (b) below, show HRO (and more generally, machine programmes) to be an excellent tool for studying *T*, it is less clear that *T*, in particular its language, has enough structure to express interesting facts about programmes. Specifically even though *T* has some non-extensional features and allows us to formulate operations which give different results when applied *to* distinct but extensionally equivalent programmes, it does not say anything about the execution *of* programmes. On a more technical level, even the model HRO is somewhat defective because HRO is introduced by Kleene's *T* predicate. This provides a good normal form for recursive *functions*, but not for *computation procedures*. As mentioned in §3b(i), Kleene lays out in an ω-order all possible computations, that is (finite) deductions in his particular equation calculus, and a computation 'from' a recursion equation say *e* consists in going through this list till one hits a deduction which really is such a computation! (Whatever else may be in doubt this is a different procedure from simply carrying out the relevant, last, deduction).

(ii) Models more directly related to *T* can be defined by means of the concept of *hereditarily computable term* introduced in [44] (where 'convertible' is used for 'computable'). The essential step is to pick out, from among the (equational) axioms, *defining equations* for constants *t* say, of the form $t = t_1$. They are used as *computation rules*, determined by the immediate reduction relation, say $\circ \rightarrow$, where $t' \circ \rightarrow t''$ means:

t'' is the result of replacing the subterm *t* of t' by t_1.

The relation $\circ\to$ is primitive recursive and finitary.[14] The terminology is justified because of the strong computability property of $\circ\to$: Every reduction sequence ρ starting with t terminates in an irreducible term $|t|_\rho$ say, and $|t|_\rho, |t|_{\rho'}$, are *unique up to congruence*; this normal form of t is denoted by $|t|$.

The *metamathematical principles* needed are

quite elementary for the proof of *congruence*, i.e., of *consistency of computation rules,*

ϵ_0-recursion for *existence* of normal forms in the sense that for $\alpha < \epsilon_0$ there is a t_α of type $0 \to 0$ such that $|t_\alpha s^n 0| = s^{fn}0$ and f is not α-recursive. (This also shows that the *transitive closure* of $\circ\to$ is not α-recursive.)[15] To be precise one needs sequence variables to *state* that *every* ρ terminates but the proof does not need strong existential axioms on ρ and can be formalized in a conservative extension of ϵ_0-arithmetic; cf. also b(i).

To describe three models of T ($\mathfrak{T}$ for 'term', $\mathfrak{N}$ for 'normal term', $\mathfrak{M}$ for 'mapping') we use $\equiv$ for congruence, t for *closed* terms, t_1 for closed terms of type σ if t is of type $\sigma \to \tau$. Let $\bar{t}$ denote the *mapping* (function): $|t_1| \mapsto |tt_1|$ if t is of type $\sigma \to \tau$, and $|t|$ if t is of type 0, and let $\bar{\underline{\circ}}$ denote *extensional equality* between such mappings. The *domains* of the models below consist of realizations of (all the) closed terms.

	$\mathfrak{T}$	$\mathfrak{N}$	$\mathfrak{M}$
t	t	$|t|$	$\bar{t}$
$t = t'$	$|t| \equiv |t'|$	$|t| \equiv |t'|$	$\bar{t} \; \bar{\underline{\circ}} \; \bar{t'}$
(possibly)			
juxtaposition	$t, t_1 \mapsto tt_1$	$|t|, |t_1| \mapsto |tt_1|$	$\bar{t}, \bar{t_1} \mapsto (\overline{tt_1})$

Thus, $=$ is, essentially, realized by identity in $\mathfrak{N}$ (but not in $\mathfrak{T}$). Usually juxtaposition is treated as a purely syntactic term combinator and not given a realization by some kind of application operation (it is not function *application* itself not even in $\mathfrak{M}$). This matter affects the recursion theoretic complexity of the description of $\mathfrak{N}$, since the set of normal terms and the con-

---

[14] The notation in [44] is different. Also, for the sake of the particular method of proof used, Tait picks a (more deterministic) subset of these computation rules. Note for reference that the converse of $\circ\to$ is not finitary since e.g. for all n, $s^n 0.0 \circ\to 0$.

[15] These questions concerning methods of proof and definition are to be distinguished from those concerning the (extensional order) *type* of this relation; cf. the digression in §2a(i) and §3a(iii) of the main text, and b(ii) below.

gruence relation are primitive recursive, but the mapping $|t|, |t_1| \mapsto |tt_1|$ is not.

The *metamathematical principles* needed to establish that these objects are models of T includes ϵ_0-recursion since the mapping: $t \mapsto |t|$ intervenes. (If, in the pedantic fashion of [22], p.202(b), the realization of a formula with variables involves a *proof*, we can take proofs denoted by derivations in T itself.) Evidently the *consistency* of T asserts more than the consistency of the computation rules (determined by the defining equations) of T, because T contains also such *inference rules* as induction and substitution which allow shortcuts; for example we derive $0 + a = a$ and infer $0 + s^n 0 = s^n 0$ by substitution without going through the computation of $0 + s^n 0$ from the defining equations for addition.

Remark. As would be expected from the formalist conception which suppresses type distinctions, both the *data* (the computation relation $\circ\rightarrow$) to be considered and our *assertions about the data* (formulated in T) are presented by similar formal procedures. But the roles are different; in the language of model theory, $\circ\rightarrow$ determines (more or less directly) the diagram of *given* models while the axioms and rules of T are *shown* to be valid for them; hence the expression 'inference' rule above. The particular models $\mathfrak{T}$, $\mathfrak{N}$, $\mathfrak{M}$ are *canonical* in the sense that T has a *name* for each element in their domains.

(iii) Leaving to (b) a discussion of the uses of (ii) at the present time, we may mention *routine* generalizations of (ii). Thus, reversing the historical order, instead of selecting defining equations from a given equational calculus, one would start with a suitable class of computation relations $\circ\rightarrow$, and look for equational calculi valid for them in the sense of a(ii); note that a(ii) makes sense even if $\circ\rightarrow$ is defined for objects which have no name in the equational calculus. ('Suitable' means that the class is large, yet satisfies lots of, equational, axioms and rules). Or we might look for manageable syntactic conditions on sets of equations which ensure strong computability (of all closed terms built up from the symbols in the equations); in other words, we ask: What are defining equations? As a corollary one would expect information on *general* extensions of classes of defining equations which preserve computability such as the familiar addition of an *equality functional E* with the reduction rules, for closed terms t, t':

$$\text{From } |t| \equiv |t'| \text{ derive } E(t, t') = 1$$
$$\neg |t| \equiv |t'| \text{ derive } E(t, t') = 0.$$

Warning. For sensible conjectures it is necessary to keep a firm grasp of the distinction between computations or conversions on the one hand and proofs on the other (even if within some limited context we may happen to have some extensional equivalence). This is well illustrated by allowing *free variables in the defining equations* of *T*. Their computability, to normal terms also possibly containing free variables, is shown as for closed terms and, consequently, $\{(t', t''): |t'| \equiv |t''|)\}$ is recursive. Inspection shows, for constants t', t'' of type $\sigma \to \tau$ and *variable* a of type σ that

$$|t'| \equiv |t''| \Leftrightarrow |t'a| \equiv |t''a| \ .$$

Also, of course

$$|t'| \equiv |t''| \Rightarrow \vdash_T t' = t'' \ .$$

But while the converse holds for *closed* t', t'', it does not hold, in general, for open terms. In fact, for primitive recursive t', t'' of type $0 \to 0$ the sets

$$\{(t', t''): \vdash_T t'a = t''a\} \text{ and } \{(t', t''): \exists n \vdash_T t's^n 0 \neq t''s^n 0\}$$

are not even recursively separable. Naturally it would be interesting to analyze explicitly the distinction between computation and inference rules involved here but in the mean time one had better respect it implicitly if one wants to work in the field at all.

Model theoretic terminology is suitable to describe the facts above. Thus *T* is *sound* for its own computation rules, that is, for closed terms t' and t''

$$(\vdash_T t' = t'') \Rightarrow |t'| \equiv |t''|$$

and *T* is *complete* since

$$(|t'| \equiv |t''| \Rightarrow \vdash_T t' = t'')$$

(the latter even for *open* t' and t'').

(b) *Reviewing the situation.* Without going into the reasons why the models HRO in a(i), $\mathfrak{T}$ or $\mathfrak{N}$ in a(ii) were introduced or why people wanted to assign ordinals to terms of *T* in (i) below, we shall consider some philosophical and mathematical uses.

(i) Gödel's intention in [7] was to use (the properties of the notion of constructive function of finite type codified in) T for a *consistency proof* of formal arithmetic. The models HRO, $\mathfrak{T}$ or $\mathfrak{N}$ are not suitable because they are, literally, defined in arithmetic terms, with this difference: HRO is even extensionally non-arithmetic, while $\mathfrak{N}$ has at least a primitive recursive domain and = (though not application). The *proof*, e.g. in [44], showing that $\mathfrak{N}$ is a model uses full arithmetic together with the reflection principle for $\forall\exists$ formulae. (It applies induction to the logically complicated computability predicate to show that the latter is primitive recursive! all terms being computable.) This is well known to be proof theoretically equivalent to ϵ_0-recursion, by Gentzen's syntactic transformations discussed in §2.

Alternatively *this* proof theoretic equivalence can now be established by the formal reduction to T in [7] together with the assignment of ordinals $< \epsilon_0$ to terms of T in [9] (or the earlier but neglected work [58]), the meta-mathematical principles used being identical. The total labour required for the two alternatives is probably much the same in the particular case of formal arithmetic here considered; see §2b and the survey [55] for advantages of the second alternative (the functional interpretation) for *extensions* of arithmetic, when an ordinal analysis is unwieldy. Naturally this route is really attractive if the relevant computation rules are found to have independent interest.

In [7] definitional equality between terms of type $\neq 0$ was mentioned but not used. Tait expressed doubts in [44] about the sense of definitional equality $t = t'$ unless all possible *arguments* of t and t' are listed. Whatever the virtues of such a listing may be, the model HRO shows clearly that those doubts are unfounded (and so do non-canonical models referred to in a(iii) if models more in the style of [44] are preferred).[16]

Finally it is perhaps worth remarking that, in accordance with footnote 6 of [19], the mechanical character of the rules treated is not essential to the notion of definitional equality; cf. a(i).

(ii) The particular mathematical uses of $\mathfrak{N}$ emphasized in [44], for example the *conservative extension* property of T^ω (that is T with decidable = and intuitionistic logic for functions of finite type) over first order arithmetic,

---

[16] This was stressed in my review of [44]. I take this opportunity to correct an error there when I say that computability of terms is also needed for $1 -$ consistency: any model of T (over the natural numbers) provides some numerical denotation for each closed term of type 0. So if $\exists xAx$ has been proved in arithmetic for primitive recursive A, by [7] there is a term t for which At is proved in T, and if $\bar{t}$ is the value of t in the given model, $A\bar{t}$ is true and hence derivable in arithmetic. Of course without computability of the numerical term t, its denotations in different models might be distinct.

can also be proved by use of HRO. More simply, as is to be expected from §2b since HRO is a more 'brutal' model.

Presumably $\mathfrak{N}$ is more efficient when the defining equations themselves (and not general models) of T are principal objects of study. But, perhaps, $\mathfrak{N}$ is still the best tool for the following result which does not mention 'defining equations':

There is a *free* model of T, that is for closed terms t_1, t_2

$$t_1 = t_2 \text{ is true in the model iff } \vdash_T t_1 = t_2$$

and this model is *recursive.*

Still taking the interest of T for granted, a more striking use of a(ii) is the connection established between (the functionals of) T and a genuine ordered structure, specifically the (well-founded) *computation relation* induced by the defining equations, read from left to right, and presumably of (order) type ϵ_0 (ω being excluded by the remark in footnote 14). As in footnote 15, ϵ_0 has here an extensional , not an auxiliary metamathematical significance. This connection between the formal principles T and ϵ_0 is certainly not apparent from HRO nor, as already discussed in §3b, from the original interpretation of T.

However, the most fruitful step at the present time is to *look* for significant stronger results. Goodman showed (in unpublished work) that the conservative extension property holds if *axioms of choice* are added to T^ω. Here $\mathfrak{N}$ does not seem to be enough; nor HRO even though one of the most interesting results of [46] asserts that it is consistent to assume that HRO satisfies axioms of choice.

Remark on conservative extension results. Technically they serve to *formulate* explicitly vaguely felt 'differences' between different metamathematical studies; for example the simple ⌐⌐ translation shows only that classical analysis is conservative over the theory of species for $\forall$ formulae, while Spector's work was used in [19] to extend this to $\forall\exists$ formulae; cf. §2a(ii). But, at the risk of sounding a bit pompous, one can say more: they establish *adequacy or: autonomy properties of languages and principles of inference*. Recall that when we justify axioms in constructive mathematics, say in arithmetic, we refer to *proofs* and *functions* (on proofs), concepts which are not mentioned in the language of formal arithmetic at all. Indeed extending the language (and principles) constitutes the most familiar method of strengthening necessarily incomplete systems of arithmetic, for example when induction is applied to predicates in a wider language. Now conservative extension results establish *adequacy with respect to the addition of certain*

principles formulated in a wider language. It may fairly be said that once we drop the doctrinaire elements of Hilbert's programme (cf. SPT, pp. 360–361) *reductive proof theory in its proper sense requires conservative extension and not merely consistency results.*

(iii) Granted an interest in *T*, (i) and (ii) contain a useful summary of its more delicate formal properties. But as formalists are never tired of saying, the interest of formal systems depends on one's 'purposes'. The only explicit-ly stated purpose of *T*, for a consistency proof of arithmetic, is not helped by the formalistic refinements of a (ii); cf. (i) above. So what purposes *are* to be served here?

Though a(i) and a(ii) mention *computation* rules it cannot be claimed that the particular rules of *T*, operating on terms of higher type, are obviously typical of the actual practice of computer science. A more promising pur-pose can, perhaps, be found if, as in [44] , we think of *T* as formulated by use of combinators, and of *combinators as providing the language for a sys-tematic theory of formal procedures*; as *set theory* provides the language for *classical* (extensional) *mathematics.* This sensible aim is sometimes obscured by 'glamorous' but far-fetched allusions to paradoxes of self reflection.

The semantics of combinators is more sophisticated than $\mathfrak{X}$ or $\mathfrak{N}$ in a(ii) where each term *t* in (the equational calculus) *T* denotes itself or at least its normal form; in short *t* denotes an object *in* the computation relation. Com-binators would express something *about* the computation relation. Or, to use familiar model theory, *combinators would have the role of logical operations.* Equations not containing combinators should correspond to atomic formulae which denote relations *in* the models considered, while logically complicated formulae express something *about* the model denoting objects in the set-theoretic *super structure* over the model (first order formulae denote sub*sets* of the domain of the model). As philosophers sometimes say: the logical operations do not 'represent' (anything *in* the model). And the introduction of combinators, or something like them, suggests itself because the *use of logically complicated formulae has turned out to be marvellously efficient for expressing notions and facts about mathematical structures which we,* that is mathematicians who have studied these structures, *really want to know*! (For example: infinite, convergent, uniformly convergent can all be ex-pressed systematically in the language of first order arithmetic with its usual interpretation).

But here the parallel ends. The *expressive power* of the model theoretic interpretation of our familiar logical languages has been *established.* We have a substantial body of informal mathematics *and* we have *Principia* which verified that this body of notions and results can be formulated in the language

of set theory (with its ordinary interpretation). In the projected theory of formal procedures, we do not have the analogue to *Principia*; in fact, we hardly have the analogue to the body of informal mathematics mentioned above. Perhaps before searching for a language (of combinators) proper to a *systematic* theory of computations, we might ask ourselves quite soberly what we really want to know about them at all.

(c) *Formalist semantics.* The impredicative starting point is the Brouwer-Heyting explanation of the logical particles. To be precise, we have two variants, one *using* logic in the explanation itself (cf. also §4(d) in connection with Kripke's models), the other *reducing* all logic to proofs of identities, that is assertions which are not logically compound; cf., e.g., [19], note 11 on p.146 concerning 'operations' and 'judgements'.

The reader is referred to Prawitz' article [37] in this volume for two examples of formalist interpretations (which satisfy Heyting's formal laws.) One, in Chapter IV of [37], is an abstract *functional* scheme using familiar λ-terms. The other, called *operational* in App. A, is more concrete, using (normalization) operations on formal deductions in familiar systems (adjoined to arbitrary Post systems). Both use logic in the metamathematical explanations, as in the first of Heyting's variants above.[17]

There is, by now, no shadow of doubt about the value for current proof theoretic studies of such formalist interpretations which, by Chapter IV of [37], relate terms-cum-computations to derivations-cum-normalizations. Even a quite rough idea of this homomorphism has turned out to be useful! In one direction, normalization procedures for deductions in the theory of species were suggested by Girard's computations of terms [56] and, in the opposite direction, Hinata's assignment of ordinals to terms [58] was suggested by one of Takeuti's ordinal analyses of deductions.

As to the initial questions (α) and (γ) of this Appendix the references above give quite a good idea of the general aims of formalist semantics and its

---

[17] To avoid (fairly common) oversights the reader should note that Heyting does not treat explicitly the notion of *logical validity*, but rather explains the meaning of the logical operations, tacitly assuming that the *atomic* formulae are interpreted and the domains of the variables are given. A definition of *logical validity* requires us to analyze these tacit assumptions, by considering all possible domains and all possible interpretations, which is done explicitly in §4c above. Evidently Heyting makes the (unstated, but not unreasonable) assumption that we know so little about these possible domains and interpretations that any proof of logical validity must be uniform in them, treat them as parameters. And by [33] existing formal theories of species are consistent with this in the sense that if $\forall X_1 ... \forall X_n \exists x A$ is derivable, so is $\forall X_1 ... \forall X_n A[x/t]$ for a term t independent of $X_1, ..., X_n$.

mathematical uses (It seems plausible that the material on T reported in (a) and (b) can be obtained as some kind of corollary if the *computation rules* of T are construed as suitable Post rules.) It remains to consider question (β) and, of course, corresponding to b(iii), the expressive power of formalist interpretations.

(i) *Logically complex (recursive) realizability.* This is a recursive 'intermediary' between the abstract notion and its formalist substitute. To avoid purely superficial but distracting blemishes, let me begin with some trivial modifications in Kleene's original realizability (for number theory); cf. also Troelstra's exposition in this volume. We write $R(e, \alpha)$ for:
    e realizes the (closed) formula α.
For *atomic* formulae α, Kleene puts $R(e, \alpha) \Leftrightarrow \alpha$, which is natural in number theory where all atomic formulae are decidable. If instead we had some Post system for generating atomic formulae we should put
    $R(e, \alpha) \Leftrightarrow e$ derives α in the system.
(Indeed, this style of modification is used when Kleene modifies the clauses of his realizability definition by additional requirements of formal derivability or validity of α in order to get suitable *derived* rules.)

Again in the interpretation of the *universal quantifier* for number theory it is natural to hold the domain fixed. When treating general Post systems as 'approximations to our knowledge about atomic sentences' (as in 1.1 of Chapter IV in [37]), approximations to our knowledge of the domain of individuals suggest expanding rather than constant domains (familiar from Kripke's models [24]).

Finally, and quite trivially, Kleene's *type-free* realizations (e can realize different formulae) may be replaced by a typed version where now a formula α is realized by the *pair*, (e, α), where, in Kleene's version, it is realized by e itself; *application* is defined by

$$\{(e', \alpha \to \beta)\}\,(e, \alpha) = (\{e'\}(e), \beta).$$

I want to go into the realizability interpretation of *implication* more carefully because its non-standard character is often (though not always: cf. note 1 on p.143 of [19]) overlooked. It is directly relevant here because it also affects the (logically complex) formalist interpretations in [37]. We are not merely verifying some given formal laws; so we have to make the *intensions of realizability more explicit*. To repeat (from §4d and elsewhere): It would be circular to use logical operations in the definition of realizability if the latter were intended to tell us their meaning. But it is perfectly sensible to use

them (if we understand them), to render the meaning of α more explicit. This is done by introducing the relation $R(e, \alpha)$, that is

    e realizes α

with (the proposition denoted by) α *equivalent* to $\exists e R(e,\alpha)$. So $\alpha \to \beta$ is rendered by

$$\exists e' R(e', \alpha) \Rightarrow \exists e'' R(e'', \beta),$$

or

$$\forall e' [R(e', \alpha) \Rightarrow \exists e'' R(e'', \beta)].$$

Despite the fact that $R(e', \alpha)$ is in general undecidable, $R(e, \alpha \to \beta)$ is taken to be

(**) $\forall e' [R(e', \alpha) \Rightarrow R(\{e\}(e'), \beta)].$

This step is certainly non-standard, for example, the equivalence between (*) and (**) is not a formal theorem of first order arithmetic (though it can be stated in arithmetic language for fixed α and β). If, as in Kleene's original formulation, $\{e\}$ is *partial*, the interpretation means that e'' depends *only* on e' and not on other information about $R(e', \alpha)$. In short, though e'' does not depend extensionally on the function $\{e'\}$, the dependence is 'too' extensional because it ignores differences between possible proofs of $R(e', \alpha)$. The non-standard character becomes even more striking if, instead of partial functions, we consider *total* realizing functions. (This applies directly to Chapter IV of [37]; and such total 'typed' variants of Kleene's realizability also satisfy Heyting's laws; for a related example worked out in the literature, cf. the modified realizability interpretation by HRO or HEO studied in [46].) Then the passage from (*) to (**) amounts to replacing

$$\forall e' [R(e', \alpha) \Rightarrow \exists e'' R(e'', \beta)] \quad \text{by} \quad \forall e' \exists e'' [R(e', \alpha) \Rightarrow R(e'', \beta)],$$

although $R(e', \alpha)$ is an undecidable predicate (of e').

    This concludes the comments *about* recursive realizability, and we turn to the *basic refinement* corresponding to that of a(i).

> Use variables for different types and write down
> *explicit closure conditions* (on their ranges) sufficient
> to provide realizations of all formally derivable theorems.

As in b(ii) we may ask for the technical use of such refinements. Clearly they are not needed for 'brutal' results such as the formal independence of

$$\forall x \, \exists y \, \forall z \, [\, T(x, x, y) \vee \neg\, T(x, x, z)]$$

for which ordinary recursive realizability is sufficient. On the other hand the first proof of independence of Markov's principle in predicate logic ([16], p.113, 3.52) relied on such a refinement.

Remark. The great number of possible variants of realizability interpretations of the logical operations (in particular of $\forall$ and $\rightarrow$ above, all satisfying Heyting's laws) seem to me to support the view expressed in the introduction to this Appendix: not the existence *of*, but the choice *between*, formalist semantics is the most fruitful issue, at least at the present time.

(ii) *Logic-free formalist semantics*, where in contrast to realizability in (i) the metamathematical explanations do not have a logical residue. One good example is in Gödel's work [7] where an even more obviously non-standard interpretation of $\rightarrow$ is used than in realizability. Gödel's original interpretation is not formalist in the sense of this Appendix although the terms used are explicitly listed, since no *formalist* meaning is given to them. Such a meaning is supplied by [44] but *this* version is not logic-free; cf. b(i) above and also §2c(iii).

A **thoroughly logic free and formalist** semantics (satisfying Heyting's laws) can presumably be obtained by formalizing appropriately the logic-free version of the Brouwer-Heyting explanation mentioned in the first paragraph of the present subsection (c). The particular use of *partial* functions by Goodman [59] does not recommend itself for this purpose because of the, apparently, essential use of the hidden logical operation involved in the concept 'is defined'.

As already mentioned in b(ii) some refinement in this direction seems to be needed for the best formal results to date (concerning the addition of axioms of choice). If this need is genuine, the extra complication introduced by having variables over *judgements* may be tolerable; 'extra' compared to the already complicated formalist refinements of recursive realizability.

Evidently logic-free (formalist) semantics have foundational interest only because the relevant *definition principles can be given an independent justification* (as in §3 or [20]).

Remark. Again the mere existence of such semantics is not problematic, by

note 32 on p.513 of [20] applied to logic free progressions (indexed by or-dinals σ). The critical point, once more in the interpretation of $\rightarrow$, is the *height restriction* ρ_σ expressing the informal requirement that, at stage σ, the step from σ' to $\rho_\sigma(\sigma')$ is known to be justified; cf. [20], p.498. Writing $P(e, \sigma; \alpha)$ for the relation:

> e, coding a derivation e_1 and a term e_2, both at level σ,
> realizes the formulae α (at σ),

$P(e, \sigma; \alpha \rightarrow \beta)$ requires e_1 to be a formal derivation at σ of

$$P(e', \sigma'; \alpha) \Rightarrow P[e_2 e', \rho_\sigma(\sigma'); \beta]$$

with variables e' and σ'. Here, and in contrast to (i), P is *decidable*.

(iii) For orientation on *purposes of formalist semantics* it is well to re-member common features of all reductionist schemes.[18] They begin with, often doctrinaire, objections to existing notions. The reductions are not uniquely determined. Closer examination shows the objections to be more dubious than the notions in question. At this stage the 'glamour' of the re-ductionist scheme fades. Nevertheless we may return to the scheme, no longer satisfied with the mere existence of a reduction, but requiring it to help us understand matters of established interest. There are two general di-rections.

Firstly we may want to know about the objects used in the reductionist scheme. Specifically, if we consider the *operational* semantics in App. A of [37] which is about *formal rules*, we ask ourselves what we want to know about them; cf. b(iii) in connection with computations. The very first question (of any semantics) is:

> What can we express?

in the given language with our interpretation. Here is a possible 'test'. In his book *Beweistheorie* [41] pp.40–46, Schütte treats *zulässige, ableitbare, and*

---

[18] To avoid misunderstanding: formalist semantics is not *explicitly* needed for Hilbe programme, that paradigm of reductive proof theory, which does not interpret logically complex formulae at all. But such interpretations are usually needed to *find* consistency proofs and, particularly, to make the syntactic transformations intelligible; for a con-vincing example see App. A of [37].

direkt ableitbare Regeln, that is rules which are conservative (also called: derived), demonstrably derived by constructive means, respectively derived in a specific uniform way (and conservative for all extensions). These notions are perhaps not central, but they are quite natural. So one could see whether the operational interpretation (of familiar languages) is able to express at least *these* notions 'about' formal operations. More generally, it seems to me we could profit here from experience with Kripke's interpretation [24]. No doubt the original intention was to give a set theoretic *reduction* of some current non-classical notions, among them the (impredicative) meaning of the logical particles introduced by Brouwer-Heyting. It turned out to have weaknesses for this analysis; cf. the view of [24]. But besides having several *unexpected* foundational uses such as those discussed in §4, this interpretation of the usual logical language was also found to have striking *expressive power*; 'striking' because the well-known (and at one time startling) expressive power of first order predicate logic was verified for its *model theoretic* interpretation. As shown (in unpublished work) by D.Gabbay and C.Smorynski, very natural *geometric* properties of trees can here be characterized by formulae ʕ simple syntactic structure. So if one wants a general theory of the kind of r. dels treated in [24] at all, the language of predicate calculus with the interpretation of [24] may fairly be said to have intrinsic (non-doctrinaire) interest. Incidentally for *this* interpretation the restriction to the usual intuitionistic propositional operations is quite artificial; what is usually described as a 'necessity' operator □ is *here* as natural a unary propositional operator as can be desired[19]. Perhaps this point should be kept in mind in work on operational semantics and the *addition of new 'logical' operators* should be envisaged.

The second, somewhat neglected, kind of use for reductionist schemes is as *an aid in the analysis of the abstract notions themselves*! For example, the natural use of formal independence results, by unintended 'elementary' models of the axioms *A* considered, belongs here. The results should not be regarded as showing some indefiniteness or other defect of the abstract notions, but as a *defect of our present analysis* (of them) codified by *A*. And

---

[19] On the other hand the description of □ as a 'necessity' operator is quite unconvinding since no serious analysis of (objective) *possibility* is even attempted: a structure is considered to be possible if it is consistent with the positive diagram of the available information. Nor is it verified that more realistic hypotheses about objective possibility can be expressed in the language of predicate logic. (The fact that the resulting formal laws are similar to *some* 'ordinary' uses of *possible* simply reflects the superficial character of these uses.) Current probability theory and statistics provide a much more sophisticated analysis of: possibility.

independence proofs *pin-point* specific problems about these notions which actually require further analysis. (Experience suggests that one cannot rely on ordinary practice to lead automatically to such problems; after all, even when strong axioms have actually been formulated such as those of set theory, it takes mathematicians a long time to learn to use them efficiently.) Evidently as long as the reductions are used as technical auxiliaries in this way they need not have intrinsic interest.

Another use of the same kind, that is a use which depends on the reductions being more *manageable* than the abstract notions (of primary interest), is this. The study of the reductions *suggests* certain (formal) relations and we then use them, for example conversion relations (in the case of operational semantics), to *analyze, not to replace abstract relations* (such as identity of proofs denoted by derivations). This requires a *separate* investigation, and, as we have seen in § 1a and § 1c, the *exact* significance of the formal relations in the literature is still in doubt. Here it must be remarked that *stability properties* (such as an homomorphism or even isomorphism between two such structures as terms and derivations; cf. Chapter IV of [37]) are not relevant. An independent adequacy condition is needed, of the kind discussed in § 1c. Of course as we have seen, the particular proposal of § 1c was too simple minded[20]. In any case stability results provide no safeguard against a *systematic oversight* (when we overlook intended conversion and computation rules for the *same* reason) nor against *confusion* (when we overlook, say, a conversion rule *because* the corresponding computation rules is not intended).

[20] (Added in proof) More sophisticated proposals might use work on the λ-calculus and the homomorphism between λ-terms and deductions. Specifically, as Mr. Barendregt has pointed out to me, completeness in the sense of Hilbert-Post (of the *relevant* λ-calculus!) would certainly by sufficient. A paradigm of a Hilbert-Post completeness result, for equations between normalizable terms, was established by Bohm [61].

Appendix II
Addenda and Corrigenda to SPT

The major additions made in the present article are not listed. A symbol {SPT n] means reference [n] of SPT, pp. 385–388; for supplements to the incomplete references cf. [20] p. 516.

Pp. 331–332, footnote 8. It has to be remembered that, for cut-free systems, *not all inconsistencies are demonstrably interdeducible,* that is, there is no proof (in the system) of

$$(\vdash A \wedge \vdash \neg A) \leftrightarrow (\vdash B \wedge \vdash \neg B) \, .$$

where the variables A and B range over all formulae. In connection with Hilbert's programme, the relevant inconsistencies are of the form that both A and $\neg A$ are derivable where A is (the translation of) a *real,* that is *numerical* assertion. Perhaps it should be called "Hilbert consistency". Also on p. 349, 1.3, "consistency" (of Takeuti's cut-free analysis) means: Hilbert consistency. Translations of numerical assertions, specifically equations between numerals, in Takeuti's system are first order formulae. To avoid misunderstanding: it is not literally true that Takeuti "observed" the result that I attribute to him, in the sense of actually *stating* it; but he observed a fact (derivations in his system of first order formulae consist only of first order formulae) which implies the result in an elementary way and he gave an elementary proof of this fact. Actually the proof establishes consistency uniformly for all first order formulae.

P. 350. The three questions are solved; cf. [19], p. 136. Question 1 has a positive answer. Questions 2 and 3 are answered by the result:
There are primitive recursive functionals F_1 and F_2 for which the following properties can be established finitistically:
(a) Suppose (the function) ψ maps a derivation with cut into a derivation without cut of the same end formula, then $F_1 \psi$ maps any derivation of $\exists x A x$, for a canonical representation A of a primitive recursive predicate, into a numerical derivation of An for some numeral n; for short, $F_1 \psi$ established 1-consistency.
(b) Conversely, if (the function) σ establishes 1-consistency, then $F_2 \sigma$ establishes the normal form theorem.
Thus Question 2 has a negative answer. Question 3(i) has a negative answer and (consequently) Question 3(ii) a positive one.

P. 351. All questions and conjectures in (2), i.e., 1. 3 to the end of (SPT §11) are settled.

P. 357, 1.1. The material of [SPT 2] has now appeared in improved form in J.Barwise, Applications of strict Π_1^1 predicates to infinitary logic, JSL 34 (1969) 409–422.

P. 357, 1. 4–6. The question has been settled positively by: J.Gregory, On a finiteness condition for infinitary logic, Dissertation, University of Maryland, 1969.

P. 367 (c) (i). $\mathcal{C}$ should be defined to be the collection of functions which are arithmetic (and not only those which are recursive) in some complete Π_1^1-set. Otherwise Π_1^0-CA *with* set parameters is not satisfied. The rest of the argument is unaffected.

P. 371, 1. 16–17 is false. If A is Π_2^1 and the Σ_3^1 sentence $(\exists X \in 2^\omega)A$ is provable in **ZF**, so is $(\exists X \in L \cap 2^\omega)A$ by SPT Note VI b(i) 1.c. But it is consistent to add to **ZF** the assumption that all constructible sets of integers are analytic [R.B.Jensen, Notices Am. Math. Soc. 15 (1968) 189, 68T-6]. It seems to be open whether there is an A in Π_2^1 (with a single free variable X) such that $A(X^*)$ is a theorem of **ZF** for some set theoretically defined X^*, but not for any analytic X^*.

P. 375, footnote 34. Friedman's construction has appeared under the title: Bar Induction and Π_1^1-CA, JSL 34 (1969) 353–362. *Correction.* Footnote 1 of this paper is perfectly true because I did make suggestions on writing the paper; only Friedman did not adopt these suggestions, and the interest of his work is not apparent. But the interest exists and, for once, happens to be best described in autobiographical terms. (i) In footnote 25 of [17] p. 139 I mentioned a model of bar induction; what I *said* was true, namely that there is such a model whose elements are of hyper degree $<0'$; what I *meant* was false, namely that *all* such functions constitute a model. (ii) I noticed the error as soon as Thomason's results on hyper degrees were announced because his results show that the model I meant does not satisfy the pairing axiom. (iii) A correction, stating (i) and (ii) explicitly, is in footnote 34 of SPT. (iv) While SPT was in print I prepared [55] where bar recursion of transfinite degree is considered in Problems 2 and 3 on p. 156. I convinced myself that if (i) could be made to work, Problem 2 would certainly have a positive solution (and hence Problem 3 a negative one) because the schema of bar recursion of trans-

finite degree has a model by socalled continuous or countable functionals of transfinite degree. Specifically, bar induction for $L_{\omega_1\omega}$ would allow us to extend the model of [55] §3 for bar recursion of finite degree by ordinary bar induction. Hence my interest, referred to in Friedman's paper, in the extension to $L_{\omega_1\omega}$. (v) I mentioned (i) – (iv) to him; he not only made sense of (i), but noticed that Problem 2 of [55] has a positive solution even with $\Pi_1^1 - CA$ instead of $\Delta_2^1 - CA$. (vi) The solution was too late to be included in [55], but I added a reference to it in SPT, at the end of footnote 34 ("added in proof").

References

[1] W.Ackermann, Grundgedanken einer typenfreien Logik, in: Essays on the foundations of mathematics (Magnes Press, Jerusalem and North-Holland, Amsterdam, 1961) 143–155.

[2] H.B.Curry and R.Feys, Combinatory logic (North-Holland, Amsterdam, 1958).

[3] S.Feferman, Systems of predicative analysis, JSL 29 (1964) 1–30.

[4] S.Feferman, Systems of predicative analysis II; Representation of ordinals, JSL 33 (1968) 193–220.

[5] S.Feferman, Predicatively reducible systems of set theory, Proc. Symp. Pure Math. 13 (AMS, Providence, 1971).

[6] S.Feferman, Ordinals and functionals in proof theory (ICM, Nice, 1970).

[7] K.Gödel, Uber einer bisher noch nicht benutzte Erweiterung des finiten Standpunktes, Dialectica 12 (1958) 280–287.

[8] A.Heyting, Die formalen Regeln der intuitionistischen Logik, S B Preuss. Akad. Wissenschaften, phys-math. Klasse (1930) 42–56.

[9] W.A.Howard, Assignment of ordinals to terms for primitive recursive functionals of finite type, in: Intuitionism and proof theory, eds. A.Kino, J.Myhill and R.E.Vesley (North-Holland, Amsterdam, 1970) 443–458.

[10] R.B.Jensen and M.E.Schroder, Mengeninduktion und Fundierungsaxiom, Arch f. math. Logik u.Grundlagen 12 (1969) 119–133.

[11] L.Kalmar, Über arithmetische Funktionen von unendlich vielen Variablen, welche an jeder Stelle nur von einer endlichen Anzahl von Variablen abhängig sind, Colloquium mathematicum 5 (1957) 1–5.

[12] S.C.Kleene, Forms of predicates in the theory of constructive ordinals, American Journal of Math. 66 (1944) 41–58; 77 (1955) 405–428.

[13] S.C.Kleene, Recursive functionals and quantifiers of finite types, Trans. A.M.S. 91 (1959) 1–52.

[14] G.Kreisel, On the interpretation of non-finitist proofs, JSL 16 (1951) 241–267.

[15] G.Kreisel, Ordinal logics and the characterization of informal concepts of proof, (ICM, Edinburgh, 1958) 289–299.

[16] G.Kreisel, Interpretation of classical analysis by means of constructive functionals of finite type, in: Constructivity in mathematics, ed. A.Heyting (North-Holland, Amsterdam, 1959) 101–128.

[17] G.Kreisel, Mathematical logic, in: Lectures on modern mathematics, vol. III, ed., Saaty (N.Y., 1965) 95–195.

[18] G.Kreisel, A survey of proof theory, JSL 33 (1968) 321–388.
[19] G.Kreisel, Church's thesis; a kind of reducibility axiom for constructive mathematics, in: Intuitionism and proof theory (North-Holland, Amsterdam, 1970) 121–150, reviewed Zentralblatt. 199 (1971) 300–301.
[20] G.Kreisel, Principles of proof and ordinals implicit is given concepts, in: Intuitionism and proof theory (North-Holland, Amsterdam, 1970) 489–516.
[21] G.Kreisel, The collected works of Gerhard Gentzen, Journal of Philosophy 68 (1971) (to appear).
[22] G.Kreisel and J.L.Krivine, Elements of mathematical logic; model theory (North-Holland, Amsterdam, 1967).
[23] G.Kreisel and A.S.Troelstra, Formal systems for some branches of intuitionistic analysis, Annals of mathematical logic 1 (1970) 229–387.
[24] S.A.Kripke, Semantical analysis of intuitionistic logic I, in: Formal systems and recursive functions, eds. J.N.Crossley and M.A.E.Dummett (North-Holland, Amsterdam, 1965) 92–130.
[25] P.Martin-Löf, Notes on constructive mathematics (Uppsala, 1970).
[26] P.Martin-Löf, this volume.
[27] P.Martin-Löf, Hauptsatz for the simple theory of types (to appear).
[28] G.E.Minc, Analog of the Herbrand's theorem for the non-prenex formulas of the constructive predicate calculus, Sem. in Math., V.A.Steklov Math. Institute, (Leningrad 4, 1969) 47–51, reviewed Zentralblatt 186 (1970) 5–6.
[29] J.Myhill, The formalization of intuitionism, in: Contemporary philosophy, ed., Klibansky (Florence, 1968) 324–341.
[30] L.Pozsgay, Liberal intuitionism as a basis for set theory, Proc. Symp. Pure Math 13 (AMS, Providence, 1971).
[31] D.Prawitz, Angående konstruktiv logik och implikationsbegreppet, in: Sju filosofiska studier tillägnade Anders Wedberg (Stockholm, 1963) 9–32, reviewed JSL 33 (1968) 605.
[32] D.Prawitz, Natural deduction, A proof theoretical study (Almquist & Wiksell, Stockholm, 1965).
[33] D.Prawitz, Some results for intuitionistic logic with second order quantifier-rules, in: Intuitionism and proof theory (North-Holland, Amsterdam, 1970) 259–269.
[34] D.Prawitz, Constructive semantics, Proc. of first Scandinavian logic symposium, Abo 1968 (Uppsala, 1970).
[35] D.Prawitz, On the proof theory of mathematical analysis, in: Logic and Value, Essays dedicated to Thorild Dahlquist on his fiftieth birthday (Uppsala, 1970).
[36] D.Prawitz, The philosophical position of proof theory, in: Contemporary philosophy in Scandinavia (Baltimore, 1970).
[37] D.Prawitz, this volume.
[38] H.Putnam, Foundations of set theory, in: Contemporary philosphy, ed. Klibansky (Florence, 1968) 275–285.
[39] B.Scarpellini, On cut elimination in intuitionistic systems of analysis, in: Intuitionism and proof theory (North-Holland, Amsterdam, 1970) 271–285.
[40] B.Scarpellini, A model for bar recursion of higher types, Compositio Math. 23 (1971) 123–153.
[41] K.Schütte, Beweistheorie (Springer, Berlin, 1960).
[42] J.R.Shoenfield, Mathematical Logic (N.Y., 1967).
[43] C.Spector, Provably recursive functionals of analysis, Recursive Function Theory, Proc. Symp. Pure Maths. 5 (1962) 1–27.

[44] W.W.Tait, Intensional interpretations of functionals of finite type I, JSL 32 (1967) 198–212, reviewed Zentralblatt 174 (1969) 12.–13.

[45] G.Takeuti, A formalization of the theory of ordinal numbers JSL 30 (1965) 295–317.

[46] A.S.Troelstra, this volume.

[47] E.Wette, Definition eines (relativ vollständigen) Systems konstruktiver Arithmetik, in: Foundations of Mathematics, eds. J.J.Bulloff, T.C.Holyoke, S.W.Hahn (Springer, Berlin, 1969) 130–195, reviewed JSL 36 (1971).

[48] M.Yasugi, Interpretations of set theory and ordinal number theory, JSL 32 (1967) 145–161.

[49] J.Zucker, Dissertation, Stanford University (1971).

[50] D.Scott, Constructive validity, in: Symposium on automatic demonstration, Springer Lecture Notes 125 (1970) 237–275.

[51] S.Feferman, Formal Theories for transfinite iterations of generalized inductive definitions and some subsystems of analysis, in: Intuitionism and proof theory (North-Holland, Amsterdam, 1970) 303–326.

[52] S.Feferman, Hereditarily replete functionals over the ordinals, in: Intuitionism and proof theory (North-Holland, Amsterdam, 1970) 289–302.

[53] W.A.Howard, Assignment of ordinals to terms for type zero bar recursive functionals (abstract) JSL 35 (1970) 354.

[54] G.Kreisel, Two notes on the foundations of set-theory, Dialectica 23 (1969) 93–114.

[55] G.Kreisel, Functions, ordinals, species, in: Logic, methodology and philosophy of science III (North-Holland, Amsterdam, 1968) 145–159, reviewed Zentralblatt 187 (1970) 265–266.

[56] J.Y.Girard, this volume.

[57] A.Levy, Principles of reflection in axiomatic set theory, FM 49 (1960) 1–10.

[58] Shigeru Hinata, Calculability of primitive recursive functionals of finite type, Science Reports of the Tokyo Kyoiku Daigaku, A, 9 (1967) 218-235.

[59] N.D. Goodman, A theory of constructions equivalent to arithmetic, in: Intuitionism and proof theory (North-Holland, Amsterdam, 1970) 101–120.

[60] H.Luckhardt, Habilitationsschrift, Philipps-Universität, Marburg, 1970.

[61] C.Böhm, Alcune proprieta delle forme β-η-normali nel λ-K-calcolo, Pubblicazioni dell'istituto per le applicazioni del calcolo, no. 696, Consiglio nazionale delle recerche, Rome, 1968.