A most artistic package of a jumble of ideas

Fernando Ferreira* Universidade de Lisboa ferferr@cii.fc.ul.pt

Abstract

In the course of ten short sections, we comment on Gödel's seminal "Dialectica" paper of fifty years ago and its aftermath. We start by suggesting that Gödel's use of functionals of finite type is yet another instance of the realistic attitude of Gödel towards mathematics and in tune with his defense of the postulation of ever increasing higher types in foundational studies. We also make some observations concerning Gödel's recasting of intuitionistic arithmetic via the "Dialectica" interpretation, discuss the extra principles that the interpretation validates, and comment on extensionality and higher order equality. The latter sections focus on the role of majorizability considerations within the "Dialectica" and related interpretations for extracting computational information from ordinary proofs in mathematics.

Ι

Kurt Gödel's realism, a stance "against the current" of his time, is now wellknown and documented. Later in life, Gödel even wrote that his realism was important for his mathematics (of course, retrospective judgements have to be approached with much caution). In a letter to Hao Wang in December 7, 1967, Gödel wrote: "This blindness (or prejudice, or whatever you may call it) of logicians is indeed surprising. But I think the explanation is not hard to find. It lies in a widespread lack, at that time, of the required epistemological attitude toward metamathematics and toward non-finitary reasoning" (see [Göd03b]). The fruitfulness of this "epistemological attitude" for Gödel's work is usually illustrated by his proof of the completeness theorem (which is of necessity nonfinitary, since it must use weak König's lemma or something equivalent) and by his definition of the constructible universe (which takes all ordinals for granted, instead of trying to "construct" the ordinals as well). These may be considered illustrations of his set-theoretic realism. I suggest a third illustration of Gödel's realistic stance in mathematics, viz the introduction of the notion of computable functional of finite type over the natural numbers in his Dialectica paper of 1958. The latter illustration has the benefit of exemplifying in a clear way Gödel's wide

^{*}This work was partially supported by ${\ensuremath{\sf CMAF}},\,{\ensuremath{\sf POCI2010/FCT}}$ and ${\ensuremath{\sf FEDER}}.$

realism in mathematics, one that goes beyond the mere set-theoretic – indeed, a conceptual realism of sorts.¹

The stated purpose of Gödel's Dialectica paper is to present a consistency proof of first-order Peano arithmetic PA by way of an extension of the finitary viewpoint according to which the requirement that constructions must be of and over "concrete" objects is relaxed to be of and over certain abstracta, namely computable functionals of finite type over the natural numbers. Gödel defines a truth-functional quantifier-free theory T whose terms are built from symbols denoting the new abstracta and whose formulas are Boolean combinations of equations between the terms. There is an evident formal parallelism between T and the theory PRA (primitive recursive arithmetic): T is, essentially, PRA extended with new terms. The *Dialectica* interpretation assigns to each formula A of the language of first-order arithmetic a formula $A_D(x,y)$ of the language of T. The technical result says that if a sentence A is derivable in Heyting arithmetic HA, then there is a closed term t of the language of T such that $A_D(t,y)$ is derivable in T.² In particular, the consistency of HA (and therefore, via the Gödel-Gentzen negative translation, that of PA) is reduced to the consistency of T because $(0 = 1)_D$ is simply 0 = 1.

The move of extending the notion of calculable function, or effective rule, to finite types is in line with Gödel's view of mathematics and, in particular of set theory, according to which the postulation of ever increasing types in set theory – via large cardinal axioms – is the royal road to the study of the universe of sets. It is significant that in a lecture of 1941 (posthumously published in [Göd41]), and that both in the *Dialectica* paper and its 1972 descendant ([Göd58] and [Göd72], respectively), Gödel suggests extending the notion of effective rule to (constructive) transfinite types. Effective rules in higher types give rise to new effective rules which compute additional functions from the natural numbers $\mathbb N$ to the natural numbers $\mathbb N$. For instance, the language of $\mathbb T$ has a closed term of type 1 (i.e., denoting a function from $\mathbb N$ to $\mathbb N$) which computes the Ackermann function (via a set of reduction rules in the technical sense of a normalization procedure). In analogy with set theory, where new types give rise to new Π_1^0 -consequences, the acceptance of effective rules in higher types gives rise to new effective rules for computing functions of type 1.

The new consistency proof of Gödel must rely only on the extended finitism and should not use intuitionistic logic. The lecture of 1941, in which Gödel already presents the interpretation of the 1958 paper, reveals a disatisfaction with the "lack of precision" (see p. 4 of [Göd41]) of the intuitionistic notion of proof and with the ordinary view of intuitionistic logic.⁴ In the lecture, Gödel presented his interpretation as a recasting of intuitionistic reasoning – in the context of number theory – via the extended finitary standpoint achieving, as it were, a replacement of the intuitionistic notion of proof by the more precise notion of computable functional. Given this background, it would be of no "epistemological significance [if] the concept of computable function used and the insights that these functions satisfy the axioms of T (...) implicitly involve intuitionistic logic or the concept of proof as used by Heyting" (see note 6 of [Göd72]).

In defining the computable functionals of finite type, Gödel relies on the primitive concept of operation, always performable⁵ (and constructively recognized as such), on given computable functionals of appropriate (lower) type. He simply takes this concept "as immediately intelligible." Within this arena, Gödel accepts particular functionals (known today as the primitive recursive functionals in the sense of Gödel) as being computable. A paradigmatic case is iteration: let g be a computable functional of type $\sigma \to \sigma$, i.e., a computable functional from computable functionals of type σ to computable functionals of type σ ; one must deem computable the functional that assigns to each natural number n and computable functional h of type σ the computable functional $g^n(h)$ of type σ . According to Gödel, "one may doubt whether we have a sufficiently clear idea of the content of this notion [computable function of finite type], but not that the axioms [of T] hold for it." I agree with this statement insofar as it says that the primitive constants of T (e.g., constants for primitive recursion in finite-types)⁶ denote evidently computable functionals. This point must be granted if a discussion on the foundational aspects of the Dialectica paper is to be made at all.⁷

It is part and parcel of Gödel's extended finitism that closed equations must be decidable. The foundational justification of the truth-functional interpretation of T lies just in this fact. But how can we decide whether t=q, even for t and q closed terms of the type 0 (the type of \mathbb{N})? William Tait showed in [Tai67] that these terms reduce mechanically to numerals (the terms normalize to numerals). Numerals are the same, or not, and this decides the equation. Nevertheless, as aptly observed by Georg Kreisel in p. 112 of [Kre87], the normalization of terms is not apparent by the inspection of the notion of computable functional of finite type.^{8,9} There is an *intolerable gap* between the obviousness of the acceptance of Gödel's effective rules in finite types (as one does with iteration) and the lack of evidence for the normalization of closed terms of type 0. In what concerns the decidability of closed equations of T, the alleged "immediate intelligibility" of the notion of computable functional does not seem to pay off. I reckon that this gap was the source of the intellectual hesitations of Gödel with respect to the project of publishing an English translation of his 1958 Dialectica paper. From the time that Paul Bernays brought the matter to Gödel's attention in September 1965, almost eight years of vicissitudes passed until the project finally faded away.

II

Even though Gödel tried to found intuitionistic logic – at least within the scope of number theory – in terms of the notion of computable functional of finite type, he did not view his recasting as being faithful to the intended meaning of intuitionistic logic. He was explicit on that: in the 1958 paper, Gödel says that "obviously, we do not claim that [the clauses of the *Dialectica* interpretation] reproduce the meaning of the logical particles introduced by Heyting and Brouwer." It is precisely this unfaithfulness that makes the *Dialectica* interpretation mathematically interesting and surprising. It is an obvious point that the more faithful you are the less surprises you get.

One of the nicest features of the *Dialectica* interpretation is that it interprets both formulas $\exists n \neg A_{qf}(n)$ and $\neg \forall n A_{qf}(n)$ in essentially the same way (for quantifier-free $A_{\rm qf}$). Their equivalence is Markov's principle. Though not intuitionistically valid, Markov's principle has a valid computational interpretation: the acceptance of the latter statement assures that an unbounded search procedure will eventually halt with a natural number n such that $\neg A_{qf}(n)$. One cannot but agree with Tait in his assessment given in the last paragraph of [Tai06]: "I would rather view Markov's principle as an example of why, if one is looking for methods of proof which automatically yield algorithms for computing a witness for existential theorems, intuitionistic logic is too narrow." Yet, the Dialectica interpretation has more surprises on this regard. Under a simple extension of the interpretation to HA^ω (the intuitionistic theory obtained from T by adjoining quantifications in all finite types) the *Dialectica* still gives essentially the same interpretation to both $\exists x \neg A_{qf}(x)$ and $\neg \forall x A_{qf}(x)$, where x may be of any finite type σ . Observe that in higher types it is certainly a misnomer to call the equivalence between $\exists x \neg A_{qf}(x)$ and $\neg \forall x A_{qf}(x)$ 'Markov's principle' since there is no search procedure available for finding a witness for $\neg A_{qf}(x)$.

The soundness theorem of the Dialectica interpretation holds for HA^ω adjoined with Markov's principle (in higher types) MP^ω , as well as with the axiom of choice AC^ω and the independence of premisses principle $\mathsf{IP}^\omega_\forall$ for universal antecedents. The principles are:

$$\begin{split} \mathsf{MP}^\omega\colon \neg\forall x A_{\mathrm{qf}}(x) &\to \exists x \neg A_{\mathrm{qf}}(x), \\ \mathsf{AC}^\omega\colon \forall x \exists y B(x,y) &\to \exists f \forall x B(x,f(x)), \\ \mathsf{IP}^\omega_\forall\colon (\forall x A_{\mathrm{qf}}(x) &\to \exists y C(y)) &\to \exists y (\forall x A_{\mathrm{qf}}(x) &\to C(y)), \end{split}$$

where x and y may be of any type, $A_{\rm qf}$ is quantifier-free and B, C are arbitrary. An extension of the soundness theorem of Gödel can the stated as follows:

Theorem. If $\mathsf{HA}^\omega + \mathsf{MP}^\omega + \mathsf{AC}^\omega + \mathsf{IP}^\omega_\forall \vdash A$, where A is a sentence, then for some closed terms t, $\mathsf{T} \vdash A_D(t,y)$.

A fortiori, if we let A^D be $\exists x \forall y A_D(x,y)$, we get $\mathsf{HA}^\omega \vdash A^D$. What is the relationship between A and A^D ? When A follows from A^D in HA^ω we obtain a conservation result. This is the case for formulas in prenex normal form. In general, we have the following characterization theorem: 10

Theorem.
$$\mathsf{HA}^\omega + \mathsf{MP}^\omega + \mathsf{AC}^\omega + \mathsf{IP}^\omega_\forall \vdash A \leftrightarrow A^D$$
.

The above three principles are called the *characteristic principles* of the *Dialectica* interpretation. The characterization theorem also ensures that we are not missing any principles besides MP^ω , AC^ω and $\mathsf{IP}^\omega_\forall$ in the statement of the extended soundness theorem. To see this suppose that we could state the soundness theorem with a further principle P. Since P is a consequence of itself, from soundness it would follow that $\mathsf{HA}^\omega \vdash \mathsf{P}^D$. By the characterization theorem, we get $\mathsf{HA}^\omega + \mathsf{MP}^\omega + \mathsf{AC}^\omega + \mathsf{IP}^\omega_\forall \vdash \mathsf{P}$. In conclusion, P is superfluous.

Gödel understood equality between higher-type functionals as intensional equality in the sense that the functionals have the same procedure of computation (definitional equality). This notion only has a precise sense within a particular framework. Gödel was well aware of this. He writes the following in a letter to Bernays in July 14, 1970: "The mathematicians will probably raise objections against [intensional equality], because contemporary mathematics is thoroughly extensional and hence no clear notions of intensions have been developed. But it is nevertheless certain that, at least within the framework of a particular language, completely precise concepts of this kind could be defined" (see [Göd03a]). No precise proposal was advanced. A good candidate is to say that two terms are definitionally equal if they have the same normal form (thus obtaining the term model of T, described in 4.3 of [Tro90]). Another interpretation is considered by Gödel in note g of his 1972 paper: to identify functionals as arguments or values of higher-type functionals with the code numbers of their Turing machines. This yields the model HRO of the hereditarily recursive operations, described by Troelstra in 4.2 of [Tro90]. But one would be wrong in mistaking HRO for Gödel's intended model of T, even under the assumption of Church's thesis. In the above cited note g, as well as in the following note h, it is clear that Gödel considers HRO as a mere interpretation, among others, of T. Anyway, given Gödel's aims, the HRO proposal does not work: as explained by Troelstra in 4.4 of [Tro90], the arithmetical complexity of the statement 'n is the code number of a functional of type σ' increases with the type of σ and runs through all the stages of the arithmetical hierarchy. The punch line is that (essentially) full HA is necessary to show that HRO is a model of T.

The evidence is that Gödel considered the notion of computable functional of finite type as a primitive notion, to be made precise and analyzed via new axiomatic insights (e.g. concerning effective rules in constructive transfinite types). In tandem, Gödel also included in T primitive signs for intensional equality $=_{\sigma}$ in each non-zero type σ .¹¹ Unaccompanied by characteristic terms E_{σ} satisfying the equivalence $E_{\sigma}(x,y) = 0 \leftrightarrow x =_{\sigma} y$, it turns out that it is formally impossible to interpret the innocent looking axiom $A \to A \wedge A$.^{12,13} This is not a serious problem. One may just as well introduce the missing terms as primitives or, more elegantly, give up for good both the intensional equality signs and the accompanying terms because, for the matters at hand, it is enough to express equality between terms t and t0 of common non-zero type t0 by the replacement scheme t1 is a primitive or t2.

As opposed to definitional equality, extensional equality is expressible in HA^ω . In type 1, $\alpha =_1 \beta$ is just $\forall n(\alpha n = \beta n)$. The corresponding axiom of extensionality is $\forall \Phi^2 \forall \alpha^1, \beta^1 (\alpha =_1 \beta \to \Phi \alpha = \Phi \beta)$. Why not stick with extensional equality, as mathematicians do? The discussion around extensionality is not idle, or a mere bizarrerie of philosophers. There are also mathematical concerns (only now are we beginning to understand that the use of extensionality in mathematical proofs may pose problems for the extraction of numerical bounds; see the discussions in section 3 of [Koh05]). It is not an option for Gödel's par-

ticular interpretation (extended to HA^ω) to have extensional equality. Already the above instance of extensionality is not interpretable. It is worthwhile to see why this is so. The argument is due to William Howard in [How73]. We introduce a variant of Howard's majorizability relations due to Mark Bezem in [Bez85] (strong majorizability). These are defined inductively on the types according to the following clauses:

(a)
$$x \leq_0^* y := x \leq y$$

(b)
$$x \leq_{\rho \to \sigma}^* y := \forall u^\rho, v^\rho (u \leq_{\rho}^* v \to xu \leq_{\sigma}^* yv \land yu \leq_{\sigma}^* yv)$$

A majorizable functional F of type $\sigma \to 0$ has an important uniformity property: given a functional z of type σ , the set of natural numbers $\{F(x): x \leq^* z\}$ is bounded. In other words, there is a uniform (or common) bound for the F(x)'s, with $x \leq^* z$. Certainly, this is not the case for all type 2 set-theoretic functionals. Notwithstanding, we have a fundamental result of Howard:

Theorem. For each closed term t of the language of T, there is a closed term q such that $HA^{\omega} \vdash t \leq^* q$.

If the above form of extensionality were interpretable then there would be a closed term \mathcal{E} of type $2 \to 1 \to 1 \to 0$ such that

$$\forall \Phi^2 \forall \alpha^1, \beta^1(\alpha(\mathcal{E}(\Phi, \alpha, \beta)) = \beta(\mathcal{E}(\Phi, \alpha, \beta)) \to \Phi\alpha = \Phi\beta)$$

is provable in HA^ω and, therefore, is true. By Howard's theorem, take q a closed term such that $\mathcal{E} \leq^* q$. Consider the natural number $n = q(1^2, 1^1, 1^1)$, where 1^2 and 1^1 are the constant functions equal to 1, of appropriate types. It immediately follows that

$$\forall \Phi \leq_2^* 1^2 \, \forall \alpha \leq 1^1 \forall \beta \leq 1^1 (\forall k \leq n (\alpha k = \beta k) \to \Phi \alpha = \Phi \beta).^{17}$$

This is obviously false for Φ defined by:

$$\Phi(\alpha) = \begin{cases} 0 & \text{if } \forall k \le n+1 \, (\alpha n = 1) \\ 1 & \text{otherwise} \end{cases}$$

The above proof is *quite general*. Basically, it only uses the fact that each closed term of the theory HA^ω has a majorant.

IV

It follows from the theorem of Howard that the majorizable functionals form a model of Gödel's T. Another model of T, defined independently by Stephen Kleene (in [Kle59] under the name of countable functionals) and Kreisel (in [Kre59]), is the model ECF^{ω} of the extensional continuous functionals. A compactness argument shows that the type 2 continuous functional are majorizable. This is no longer the case for type 3 functionals. It is well known that there is a FAN functional \mathcal{F}^3 in ECF^{ω} such that

$$\forall \Phi^2 \forall \alpha \leq_1 1 \forall \beta \leq_1 1 \ (\forall k \leq \mathcal{F}(\Phi)(\alpha k = \beta k) \to \Phi \alpha = \Phi \beta).$$

In other words, $\mathcal{F}(\Phi)$ witnesses the uniform continuity of Φ restricted to the Cantor space, i.e., $2^{\mathbb{N}}$. However, the set $\{\mathcal{F}(\Phi): \Phi \leq_2^* 1^2\}$ is clearly unbounded. Therefore, \mathcal{F} is not majorizable in ECF^{ω} .

The point of this little discussion was to show that continuous functionals need not be majorizable. As we will see, it is the latter notion – not continuity – that plays the central role in the extraction of uniform bounds from mathematical proofs (cf. the discussions regarding the monotone functional interpretation in V and VI) and, via the bounded functional interpretation (cf. VII), in injecting uniformities into mathematics. The remainder of this article will concentrate on the role of majorizability within the Dialectica and related interpretations for extracting computational information from ordinary proofs in mathematics.

V

In 1993 (see [Koh93]), Ulrich Kohlenbach introduced a slight modification of Gödel's functional interpretation, dubbed the monotone functional interpretation. It is based on the same assignment of formulas but it weakens the conclusion of the soundness theorem. Instead of the existence of a closed term t such that $\mathsf{HA}^\omega \vdash \forall y A_D(t,y)$, the new conclusion only demands that t is a majorant of a witness, i.e., that $\mathsf{HA}^\omega \vdash \exists x \leq^* t \, \forall y A_D(x,y)$. What is accomplished by weakening the conclusion of Gödel's theorem?

In his first published paper [Kre51], Kreisel stressed that (for sensible theories) the addition of true (e.g., all true) Π_1^0 -sentences does not affect the class of provably recursive functions. In the context of the Dialectica interpretation, this observation cashes in as follows: universal postulates (i.e., of the form $\forall x A_{qf}(x)$, with x of any type and A_{qf} quantifier-free) can be adjoined to the theory $HA^{\omega} + MP^{\omega} + AC^{\omega} + IP^{\omega}_{\forall}$ in the statement of the extended soundness theorem provided that they are adjoined to the verifying theory HA^{ω} . The reason for this is exceedingly simple, viz that a universal sentence is essentially its own interpretation under the *Dialectica*. NB if the universal postulates are true then the verification of the behaviour of the extracted terms follows from true postulates and, therefore, is sound. In short: to extract correct computational information (in the form of a term) via the *Dialectica* interpretation, one can adjoin true universal sentences even though the evidence for their truth is not manifested in the theory itself. For all we know, the evidence could have been gotten via esoteric transfinite means in set theory. It is important to distinguish between founding the truth of a statement and using the statement in proofs.

The monotone functional interpretation permits the generalization of Kreisel's observation to a wider class of sentences, not necessarily universal. Let us see how this works for a toy – but illuminating – example, the *lesser limited principle* of omniscience LLPO with a numerical parameter z:

$$\forall k \forall r (A_{\mathrm{qf}}(z,k) \vee B_{\mathrm{qf}}(z,r)) \rightarrow \forall k A_{\mathrm{qf}}(z,k) \vee \forall r B_{\mathrm{qf}}(z,r),$$

where $A_{\rm qf}$ and $B_{\rm qf}$ are quantifier-free and k and r are numerical variables. This principle is not intuitionistically acceptable and does not have a *Dialectica* interpretation.²⁰ Nevertheless, it has a monotone functional interpretation (in a suitable verifying theory). A simple computation shows that one must find closed terms t and q of types $0 \to 0$ and $0 \to 0 \to 0$, respectively, such that

$$(\star) \ \exists n \leq t \exists f, g \leq q \forall z \forall k, r \big(A_{\mathrm{qf}}(z, fzkr) \lor B_{\mathrm{qf}}(z, gzkr) \to (nz = 0 \to A_{\mathrm{qf}}(z, k)) \land (nz \neq 0 \to B_{\mathrm{qf}}(z, r)) \big).$$

It turns out that $t := \lambda z.1$ and $q := \lambda z\lambda k, r. \max(k, r)$ do the job. The proof is simple, though not completely obvious. One considers the modified predicates

$$\overline{A}(z,k) := \forall u \le k A_{\mathrm{qf}}(z,u) \vee \exists u, v \le k (\neg A_{\mathrm{qf}}(z,u) \wedge \neg B_{\mathrm{qf}}(z,v)) \text{ and}$$

$$\overline{B}(z,r) := \forall v \le r B_{\mathrm{qf}}(z,v) \vee \exists u, v \le r (\neg A_{\mathrm{qf}}(z,u) \wedge \neg B_{\mathrm{qf}}(z,v)),$$

and verifies that, for each $z, \forall k, r(\overline{A}(z,k) \vee \overline{B}(z,r))$. By LLPO, one gets, for each $z, \forall k \overline{A}(z,k) \vee \forall r \overline{B}(z,r)$. The function n is chosen to be 0 or 1 according to whether the first or second leg of the disjunction holds. At this juncture, we draw attention to the fact that a bit of choice is used. It is now straightforward to finish the verification of (\star) .²¹

In general, the monotone functional interpretation permits the generalization of Kreisel's observation to sentences with the syntactical form $\exists u \leq r \forall v A_{\rm qf}(u,v)$, where $A_{\rm qf}$ is quantifier-free, r is a closed term and u,v may be of any type. In other words, such sentences can be adjoined to $\mathsf{HA}^\omega + \mathsf{MP}^\omega + \mathsf{AC}^\omega + \mathsf{IP}^\omega_\forall$ in the statement of the extended soundness theorem for the monotone functional interpretation provided that they are also adjoined to the verifying theory HA^ω . Some well-known non-constructive principles can be put in this form (weak König's lemma is the pre-eminent example), whereby the analysis of proofs using these principles becomes possible. There is a small twist in this: typically the principles studied take the form $\exists u \leq rz \forall v A_{\rm qf}(u,v,z)$, with parameters z. In this situation, one must consider instead their parameter-free uniform versions $\exists u \leq r \forall v, z A_{\rm qf}(uz,v,z)$ — which have the right syntactical form — and the verification takes place with the strengthened versions. Observe that if the original principle is true, then so is its uniformization (which is obtained with a bit of choice) and, therefore, computationally correct information is still obtained.

VI

There are two other benefits of the monotone functional interpretation that I want to briefly address.

The bounds that one gets from the soundness theorem of the monotone functional interpretation enjoy uniformity properties regarding certain values. The canonical example is this: if $\forall \alpha^1 \forall \beta \leq_1 s\alpha \exists n^0 A(\alpha,\beta,n)$, where s is a closed term and A is arbitrary, is a consequence of $\mathsf{HA}^\omega + \mathsf{MP}^\omega + \mathsf{AC}^\omega + \mathsf{IP}^\omega_\forall$ plus some (true) statements of the form discussed in Part V, then it is possible to construct a closed term t such that $\forall \alpha \forall \beta \leq_1 s\alpha \exists n \leq_0 t\alpha A(\alpha,\beta,n)$. The point is that n

is bounded by $t\alpha$ and that this value is *independent* of β . This is a very simple, yet crucial, observation. Due to the presence of Markov's principle, the above example (with a suitaby reformulated base theory) also holds within classical logic (via the Gödel-Gentzen negative translation) provided that A is quantifierfree.²³ It must be noted that, in aplications, compact spaces can be coded by elements which are majorized (the Cantor space is a trivial example), and this entails the existence of bounds which are *uniform* with respect to the values in the compact spaces. Many striking uniform bounds in the field of numerical functional analysis have been obtained in the recent past by Kohlenbach and his co-workers using these ideas (see [KO03] and [Koh08]). Of course, the *a priori* theoretical knowledge that such bounds must exist is instrumental in the *applied work*. The recent introduction of new base types by Kohlenbach (normed, metric and other spaces) even allows the obtainment of new uniform bounds from metrically bounded, *not necessarily compact*, spaces.²⁴ These are exciting new developments (see [Koh05] and [Koh08]).

Kohlenbach's analysis of proofs can be effected even if proofs use principles that go well beyond Peano arithmetic. Proofs which use full second-order arithmetic can be unwound, and uniform bounds can still be obtained. This is made possible by Clifford Spector's 1962 deep generalization of Gödel's interpretation to second-order classical arithmetic using bar-recursive functionals (see [Spe62]). NB the reason for the existence of uniform bounds rests, ultimately, on the fact that Spector's bar-recursive functionals are majorizable. One must emphasize again that the important issue here is the a priori knowledge that certain uniform bounds must exist, even in the presence of full second-order arithmetic! (The theoretical bounds are staggering, but sensible bounds are usually obtained in concrete applications.)

The applied work of obtaining uniform bounds from mathematical proofs via the monotone functional interpretation has acquired a trait close to systematization (via the monotone functional interpretation) and goes by the name of proof mining.²⁵ It has been a rather successful program. A further reason for this success is that the non-constructive principles (discussed above) can be treated as "black boxes" in the actual study of concrete mathematical proofs, and this fact facilitates their analyses enormously. Moreover, the analyses by means of the monotone functional interpretation are modular and one does not need to work with fully formalized proofs, relying instead on logico-mathematical experience. If the success of proof mining carries over to more fields of mathematics, it is fair to say that this "applied program" is the heir presumptive to the program of "unwinding" of proofs launched by Kreisel in the fifties. Even though the analysis of some specific mathematical proofs was suggested and some were actually carried out (sometimes in very ingenious and illuminating ways), the "unwindings" were scarce, isolated, and the original program fell rather short of its promising beginnings.²⁶

VII

In the monotone functional interpretation, the bounds (hence, the uniformities) only show up at the end of the interpretation, in the *conclusion* of the soundness

theorem. In a sense, the monotone functional interpretation only scratches the surface. A new functional interpretation – dubbed bounded functional interpretation – injects uniformities all the way in, via a new assignment of formulas (see [FO05]). Many of the basic theoretic results of Kohlenbach have very perspicuous proofs using the new interpretation. The nature of the injected uniformities can be best described via the characteristic principles of the novel interpretation. These principles prove LLPO and weak König's lemma and refute extensionality and the limited principle of omniscience LPO. The latter is:

$$\forall \alpha (\forall n(\alpha n = 0) \lor \exists n(\alpha n \neq 0)).^{27}$$

The last couple of examples make it clear that the characteristic principles are not set-theoretically sound. The bounded functional interpretation does inject uniformities which are absent in the universe of sets and which are incompatible with it. Nevertheless, the soundness theorem guarantees that provable sentences of low complexity (e.g., Π_2^0 -sentences) are not affected by the injection, and this constitutes a tool for proving conservation results and for extracting correct (uniform) bounds.

I want to discuss two characteristic principles in the sequel. In the meantime, we must introduce a notion of *intensional* majorizability (in the sense that the majorizability relations are rule-governed). The language of HA^ω is supplemented by primitive binary relation symbols \leq_{σ} (one for each type σ) and by corresponding bounded quantifiers. The *majorizability* relations are governed by the axioms

$$x \leq_0 y \leftrightarrow x \leq y;$$

$$x \leq_{\rho \to \sigma} y \to \forall u \leq_\rho v (xu \leq_\sigma yv \land yu \leq_\sigma yv).$$

Note that we do *not* have the biconditional above (otherwise, we would fall into the Howard/Bezem extensional notion of majorizability). Instead, we have *rules*

$$\frac{A_{\mathrm{bd}} \wedge u \unlhd v \to su \unlhd tv \wedge tu \unlhd tv}{A_{\mathrm{bd}} \to s \unlhd t}$$

where s and t are terms, $A_{\rm bd}$ is a bounded formula and u and v are variables which do not occur in the conclusion. The only quantifiers in a bounded formula are the bounded quantifiers, and these are regulated by the schemata

$$\forall x \le t A(x) \leftrightarrow \forall x (x \le t \to A(x))$$
 and $\exists x \le t A(x) \leftrightarrow \exists x (x \le t \land A(x)).$

Having rules, instead of axioms, is crucial. In the presence of the characteristic principles, the *extensional* relations lead to inconsistency.²⁸ The first characteristic principle of the bounded functional interpretation that I discuss is the *intensional collection scheme*:

$$\forall x \le c \exists y \ A(x,y) \to \exists b \forall x \le c \exists y \le b \ A(x,y),$$

where b, c are of any type and A is an arbitrary formula. The principle says that if one can find a witness y so that A(x, y), for each x with $x \le c$, then there is a uniform (or common) bound for these witnesses. Notice the formal similarities of this principle with L. E. J. Brouwer's fan theorem (for c of type 1 and y of type 0). The collection scheme subsumes the so-called uniform boundedness principles introduced by Kohlenbach in [Koh96]. 30,31

The other principle is the intensional bounded contra-collection scheme:

$$\forall b \exists x \le c \forall y \le b \, A_{\mathrm{bd}}(x,y) \to \exists x \le c \forall y \, A_{\mathrm{bd}}(x,y),$$

where b, c are of any type and $A_{\rm bd}$ is a bounded formula. Of course, this principle is classically equivalent to the previous one restricted to bounded matrices. The point is that we are in an intuitionistic setting. It is not too difficult to show that weak König's lemma follows from this principle. We can give a Hilbertian reading of this principle. It permits the conclusion of the existence of an element x (with $x \leq c$) such that $\forall y A_{\rm bd}(x,y)$ from the weaker statement that such x's only exist locally, in the sense that for each b there exists x (with $x \leq c$) such that $\forall y \leq b A_{\rm bd}(x,y)$. We may regard such an x as an ideal element that works uniformly for every y and whose postulation does not affect real consequences (because of the soundness theorem of the bounded functional interpretation).

VIII

The emphasis has been on intuitionistic systems but, as already noticed, after applying the negative translation it is also possible to deal with systems of classical arithmetic and analysis. Alternatively, one can proceed directly, as Joseph Shoenfield does for Peano arithmetic in his well-known textbook [Sho67]. In the classical setting, the only characteristic principle for the *Dialectica* is the quantifier-free axiom of choice (in all types).³²

A similar situation happens with the bounded functional interpretation. An interpretation that directly injects uniformities into Peano arithmetic was recently defined in [Fer07]. Three characteristic principles (which embody the injected uniformities) are necessary in this case. One is an unsound form of choice for bounded matrices which, nevertheless, includes the sound $AC_{qf}^{1,0}$ (this is quantifier-free choice for x of type 1 and y of type 0; see the statement in Part II for the notation). Another is the intensional bounded contra-collection principle. The remainder is the (unsound) intensional majorizability scheme: $\forall x \exists y (x \leq y)$.

We believe that the uniformities embodied by the above three characteristic principles can also be injected into full second-order arithmetic, by way of Spector's bar recursive functionals.³³ What about other forms of comprehension? How far can we go? There is definitely a limit for the insertion of uniformities. For instance, simple forms of comprehension for type 1 functionals already cannot be present in the theories. Specifically, there cannot be a type 2 functional E such that:

$$\forall \alpha^1 (\mathbf{E}(\alpha) = 0 \leftrightarrow \exists n (\alpha n = 0)).$$

Otherwise, using $\mathsf{AC}^{1,0}_{\mathrm{qf}}$, it is possible to define a functional μ of type 2 satisfying $\forall \alpha \, (\exists n (\alpha n = 0) \to \alpha (\mu \alpha) = 0)$. By the intensional majorizability scheme, we get $\exists z \, (\mu \leq_2 z)$. It is easy to see that this leads to a contradiction.

IX

I will end with a speculative note that strides a familiar line already discussed. Mathematicians are very liberal (in the sense of not caring) in their use of induction (and comprehension). They are oblivious to the complexity of the statements they are inducting over. Logicians, on the other hand, are very sensitive to issues of definability and know that induction (together with comprehension) is the main reason for the advent of fast growing bounds. Nevertheless, as a matter of common mathematical experience, really fast growing functions almost never show up in ordinary mathematics. This is a puzzling phenomenon. I want to point that certain forms of induction are tame in this respect, namely induction for intensional bounded formulas. In these cases, induction takes the form $A_{\rm bd}(0) \land \forall n < m(A_{\rm bd}(n) \to A_{\rm bd}(n+1)) \to A_{\rm bd}(m)$, with $A_{\rm bd}$ a bounded intensional formula. Statements like this are dealt by the bounded functional interpretation effortlessly, with no need of recursors. They are basically selfinterpretable. To what extent can inductions in ordinary mathematics be put in this form? In other words, to what extent are inductions in ordinary mathematics tame?

The use of *tame* forms of induction is a particular case of using lemmata which have trivial bounded functional interpretations (and which are true after *flattening*, that is, after interpreting the intensional relation signs by the extensional Howard/Bezem majorizability relations). The universal closures of bounded intensional formulas have such trivial interpretations. Can lemmata of this kind formulate statements with mathematically interesting consequences?^{34,35}

X

The title of this article is taken from p. 110 of Kreisel's "Gödel's excursions into intuitionistic logic" (cf. [Kre87]). Gödel's last published paper lacks the glamour and impact of his most well-known results but this most artistic package of a jumble of ideas has aged well in fifty years and continues to be a source of work and ideas. The merits of Gödel's interpretation have been rather positive but, probably, we still do not have a full grasp of the potentialities of the Dialectica and related interpretations. Even though we have insisted on the mathematical and computational benefits of Gödel's interpretation, we also believe that the final word on its foundational merits has not yet been said. I gave a very personal report on Gödel's interpretation and its aftermath, a jumble of lines on a jumble of ideas, unashamedly partial and opinionated.

I would like to thank Thomas Strahm for the invitation to participate in this commemorating issue of Gödel's "Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes," published half a century ago in this very journal. It was an honor and a great pleasure to contribute.

Notes

¹In a letter to Paul Bernays in January 12, 1975, Gödel made an uncharacteristic callous statement regarding these matters. After showing agreement with a "cautiously platonistic point of view" endorsed by his friend, Gödel goes on to say that "to me a platonism of this kind (also with respect to concepts) seems to be obvious and its rejection to border on feeble-mindness." See [Göd03a].

 2 I am assuming that the reader is familiar with Gödel's *Dialectica* interpretation and I am taking some notational liberties. For instance, x and y stand for tuples of variables, possibly empty. Likewise, t stands for a tuple of terms (of the same arity and with the same types as the tuple x). The reader should keep in mind these liberties in the remainder of the article. Anyway, I suggest Anne Troelstra's excellent introduction in [Tro90] to Gödel's 1958 paper and to its modification of 1972 (unpublished in Gödel's lifetime) for interesting discussions and, of course, the proper technical definitions and pointers to the relevant material. An alternative (and more inclusive) option is [AF98].

³This feature is usually described by saying that T is *impredicative*. It is better described as a lack of *purity of methods* within the type-hierarchy of computable functionals.

⁴The intuitionistic meaning of the logical particles is explained in terms of what constitutes a proof of a logical compound statement in terms of the constituent statements. For instance, a proof of $A \to B$ is a construction which, when applied to a proof of A, yields a proof of B.

⁵In the 1972 paper there is an emphasis on the termination of the computations for all the inputs. Cf. the italics of the word 'any' on page 275 of [Göd72].

⁶Under a suitable setting, William Tait showed in [Tai06] that iteration and primitive recursion are interchangeable as primitive notions in finite types.

⁷This is the suggestion of Bernays in a letter to Gödel in March 16, 1972: "one must (...) use the assumption that if n is a constructible numeral and if, furthermore, a process is intuitively described which, from given numerals, again provides a numeral, then the n-fold iteration of that process can be carried out. A corresponding assumption must also be employed for the recursive definitions of functionals." See [Göd03a].

⁸Kreisel goes on saying that "in *sharp contrast to sets* where, for example, Zermelo's axioms are verified on sight for all limit ordinals from a description of segments of the cumulative hierarchy" (original italics).

⁹Tait in his normalization proof needs the full apparatus of HA in order to show that all concrete closed terms normalize.

¹⁰The extended soundness theorem is implicit in [Kre59], and the next theorem is due to Mariko Yasugi [Yas63].

 11 A curious, if misguided proposal, would be to propose a notion of *intensional* equality for type 0 as well. In this case, the problem of deciding equations between closed terms of type 0 becomes solvable by fiat, and the truth-functionality of (this modified version of) Gödel's T is immediate. Note, however, that the *Dialectica* interprets the quantifier-free part of HA (essentially) by itself and, therefore, the quantifier-free theorems of HA must be theorems of T. But HA is about numbers, whereas the modified T would be (in type 0) about *procedures* for representing numbers. E.g., x + y = y + x fails on the latter interpretation.

¹²This was pointed to Gödel by Justus Diller in 1970. For the details and the possible oversight of Gödel, see section 3.3 of [Tro90].

¹³This axiom is a contraction principle. After the work of Jean-Yves Girard on Linear Logic, contraction cannot be regarded innocent anylonger.

¹⁴This way of dealing with equality was suggested by Gödel himself in note i4 of his 1972 paper and elaborated by Troelstra in [Tro90]. An alternative is to consider a suitable *rule* of extensionality: this is Clifford Spector's option in [Spe62]. In this paper, we opt for the Gödel/Troelstra treatment of equality.

¹⁵The superscripts denote the types of the variables. We assume familiarity with this notation.

 16 The full set-theoretical model interprets the type 0 domain as \mathbb{N} , and the domain of type $\sigma \to \tau$ as constituted by all the set-theoretic functions from the domain of σ to the domain of τ . When we call a sentence of the language of HA^{ω} true or false, correct or incorrect, sound or unsound, we always mean true or false with respect to the full set-theoretical model.

 $^{17}\alpha \leq \beta$ abbreviates $\forall k(\alpha k \leq \beta k)$. Clearly, $\alpha \leq 1^1$ iff $\alpha \leq_1^* 1^1$. In general, for x,y of type $\sigma \to \rho$, $x \leq y$ abbreviates $\forall z^{\sigma}(xz \leq_{\rho} yz)$.

¹⁸For the definition of this model, see the encyclopedic [Tro73].

 $^{19}\mathrm{I}$ thank Dag Normann for this observation.

 20 It is a good exercise to show that a Dialectica interpretation of LLPO would entail the recursive separability of r.e. sets. The clause for disjunction is $(A \lor B)_D(n,x,x',y,y') := (n=0 \to A_D(x,y)) \land (n \neq 0 \to B_D(x',y'))$, with an explicit flag n deciding which way to fork. Were the clause, instead, $A_D(x,y) \lor A_D(x',y')$ then LLPO would have a trivial interpretation. However, as Gödel observes in note 1 of his 1972 paper, the Dialectica version is needed to prove the soundness of the interpretation of the inference $A \to C$, $B \to C \Rightarrow A \lor B \to C$. The alternative interpretation is not sound for this rule of inference since (again) it would entail the recursive separability of r.e. sets. As a matter of curiosity, Gödel made a faux pas in his unpublished lecture of 1941 when he wrote the wrong clause for disjunction. (On finding errors in Gödel, see the last paragraph of [Kre87].)

²¹This analysis of LLPO is related to, but not quite the same as the one in [Koh01].

²²The principle LLPO can be put in this form, since it is equivalent to (\star) within the theory $HA^{\omega} + MP^{\omega} + AC^{\omega} + IP^{\omega}_{\forall}$. Sometimes, the difficult part is to show that a given principle can be put in this form.

be put in this form.
²³The base theory must be PA^ω (the classical version of HA^ω) together with the axiom of choice restricted to quantifier-free formulas. Cf. VIII.

²⁴Interestingly, Gödel suggests extending his functional interpretation to branches of mathematics other than arithmetic in p. 18 of his 1941 lecture.

 25 The name is due to Dana Scott.

²⁶Solomon Feferman gives a similar assessment of Kreisel's "unwinding" program in [Fef96].

²⁷On p. 26 of his 1941 lecture, Gödel says that with the help of the *Dialectica* interpretation it can be shown that LPO is independent from intuitionistic logic. This can be seen, for instance, by interpreting T by the hereditarily recursive operations (see [Tro95] for details). The new bounded functional interpretation yields this independence result *directly* because it refutes LPO. Other independence results can be obtained in this manner.

 $^{28} \rm{The}$ soundness theorem of the bounded functional interpretation guarantees that the characteristic principles together with the rule-governed notion of majorizability are consistent (modulo \rm{HA}^{ω}). The inconsistency mentioned in the text makes it hard to see how such a consistency result can be obtained by model-theoretic methods.

²⁹For a nice introduction to intuitionism and intuitionistic mathematics, see [vA04].

 30 These principles usually incorporate a form of choice. We ignore this feature, for the sake of perspicuity.

 31 Kohlenbach's uniform boundedness principles use the extensional form of majorizability, but always accompanied with some caveat concerning extensionality. Take, for instance, the following form of uniform boundedness: $\forall \alpha \leq_1 1 \exists n^0 A(\alpha,n) \to \exists m \forall \alpha \leq 1 \exists n \leq m A(\alpha,n)$, and suppose that A is extensional with respect to α , that is, $\forall \alpha, \beta(\alpha =_1 \beta \land A(\alpha,n) \to A(\beta,n))$. Suppose that $\forall \alpha \leq_1 1 \exists n A(\alpha,n)$. A fortiori, $\forall \alpha \leq_1 1 \exists n A(\alpha,n)$ because $\forall \alpha(\alpha \leq_1 1 \to \alpha \leq_1 1)$. By the intensional collection scheme, there is m such that $\forall \alpha \leq_1 1 \exists n \leq m A(\alpha,n)$. It can be shown that $\forall \alpha (\min_1(\alpha,1) \leq_1 1)$, where \min_1 is the minimum function defined pointwise. Therefore, $\forall \alpha \exists n \leq m A(\min_1(\alpha,1),n)$. Now, take $\alpha \leq_1 1$. Clearly, $\min_1(\alpha,1) =_1 \alpha$. By extensionality, we get $\exists n \leq m A(\alpha,n)$.

³²This is essentially a result of Kreisel in [Kre59].

³³At the time of this writing, this has not yet been verified.

³⁴In a sense, the answer to this question is a trivial 'yes' because the statements considered by Kohlenbach (see the end of Part V) can be dealt by lemmata of this kind. The question is really meant for mathematical statements beyond those.

 35 Consider the following: we may assume that our language includes an intensional order relation $\unlhd_{\mathbb{R}}$ infixing between real numbers. The relationship between this intensional relation and the ordinary extensional (set-theoretic) one is rather close: $x<_{\mathbb{R}}y\to x\unlhd_{\mathbb{R}}y$ and $x\unlhd_{\mathbb{R}}y\to x\subseteq_{\mathbb{R}}y$. (In applications, Kohlenbach and his co-workers often switch from $<_{\mathbb{R}}$ to $\leq_{\mathbb{R}}$ and vice-versa in order to have the appropriate logical form, but this has limitations vis-à-vis the use of an intensional order.) Maybe this fact, or similar ones, are relevant for finding important lemmata as described above.

References

- [AF98] J. Avigad and S. Feferman. Gödel's functional ("Dialectica") interpretation. In S. R. Buss, editor, Handbook of proof theory, volume 137 of Studies in Logic and the Foundations of Mathematics, pages 337–405. North Holland, Amsterdam, 1998.
- [Bez85] M. Bezem. Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals. *The Journal of Symbolic Logic*, 50:652–660, 1985.
- [Fef96] S. Feferman. Kreisel's "unwinding" program. In P. Odifreddi, editor, Kreiseliana: about and around Georg Kreisel, pages 247–273. A. K. Peters, Wellesley Massachusetts, 1996.
- [Fer07] F. Ferreira. Injecting uniformities into Peano arithmetic. Manuscript, 15 pages, 2007.
- [FO05] F. Ferreira and P. Oliva. Bounded functional interpretation. *Annals of Pure and Applied Logic*, 135:73–112, 2005.
- [Göd41] K. Gödel. In what sense is intuitionistic logic constructive?, 1941. Published in [Göd95], pages 189–200.
- [Göd58] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958. Reprinted with an English translation in [Göd90].
- [Göd72] K. Gödel. On an extension of finitary mathematics which has not yet been used, 1972. Published in [Göd90], pages 271–280. Revised version of [Göd58].
- [Göd90] K. Gödel. Collected Works, Vol. II. S. Feferman et al., eds. Oxford University Press, Oxford, 1990.
- [Göd95] K. Gödel. Collected Works, Vol. III. S. Feferman et al., eds. Oxford University Press, Oxford, 1995.
- [Göd03a] K. Gödel. Collected Works, Vol. IV. S. Feferman et al., eds. Oxford University Press, Oxford, 2003.
- [Göd03b] K. Gödel. Collected Works, Vol. V. S. Feferman et al., eds. Oxford University Press, Oxford, 2003.
- [How73] W. A. Howard. Hereditarily majorizable functionals of finite type, 1973. Published in [Tro73], pages 454–461.

- [Kle59] S. C. Kleene. Countable functionals. In A. Heyting, editor, *Constructivity in Mathematics*, pages 81–100. North Holland, Amsterdam, 1959.
- [KO03] U. Kohlenbach and P. Oliva. Proof mining: a systematic way of analysing proofs in mathematics. *Proceedings of the Steklov Institute of Mathematics*, 242:136–164, 2003.
- [Koh93] U. Kohlenbach. Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallée Poussin's proof for Chebycheff approximation. Annals of Pure and Applied Logic, 64:27–94, 1993.
- [Koh96] U. Kohlenbach. Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals. *Archive for Mathematical Logic*, 36:31–71, 1996.
- [Koh01] U. Kohlenbach. Intuitionistic choice and restricted classical logic. Mathematical Logic Quarterly, 47:455–460, 2001.
- [Koh05] U. Kohlenbach. Some logical metatheorems with applications in functional analysis. Transactions of the American Mathematical Society, 357:89–128, 2005.
- [Koh08] U. Kohlenbach. Applied Proof Theory: Proof Interpretations and their Use in Mathematics. Springer Monographs in Mathematics. Springer, Berlin, to appear in 2008.
- [Kre51] G. Kreisel. On the interpretation of non-finitist proofs, part I. *The Journal of Symbolic Logic*, 16:241–267, 1951.
- [Kre59] G. Kreisel. Interpretation of analysis by means of constructive functionals of finite types. In A. Heyting, editor, *Constructivity in Mathematics*, pages 101–128. North Holland, Amsterdam, 1959.
- [Kre87] G. Kreisel. Gödel's excursions into intuitionistic logic. In Gödel Remembered, pages 65–186. Bibliopolis, Napoli, 1987.
- [Sho67] J. R. Shoenfield. Mathematical Logic. Addison-Wesley Publishing Company, 1967. Republished in 2001 by AK Peters.
- [Spe62] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles in current intuitionistic mathematics. In F. D. E. Dekker, editor, *Recursive Function Theory:* Proc. Symposia in Pure Mathematics, volume 5, pages 1–27. American Mathematical Society, Providence, Rhode Island, 1962.
- [Tai67] W. Tait. Intentional interpretations of functionals of finite type I. *The Journal of Symbolic Logic*, 32:198–212, 1967.
- [Tai06] W. Tait. Gödel's interpretation of intuitionism. *Philosophia Mathematica* (III), 14:208–228, 2006.

- [Tro73] A. S. Troelstra (ed.). Metamathematical Investigation of Intuitionistic Arithmetic and Analysis, volume 344 of Lecture Notes in Mathematics. Springer, Berlin, 1973.
- [Tro90] A. S. Troelstra. *Introductory note to* [Göd58] and [Göd72], 1990. Published in [Göd90], pages 217–241.
- [Tro95] A. S. Troelstra. *Introductory note to* [Göd41], 1995. Published in [Göd95], pages 186–189.
- [vA04] M. van Atten. On Brouwer. Wadsworth Philosophers Series. Thompson Wadsworth, 2004.
- [Yas63] M. Yasugi. Intuitionistic analysis and Gödel's interpretation. *Journal* of the Mathematical Society of Japan, 15:101–112, 1963.