PROOF INTERPRETATIONS AND MAJORIZABILITY

FERNANDO FERREIRA

Abstract. In the last fifteen years, the traditional proof interpretations of modified realizability and functional (*dialectica*) interpretation in finite-type arithmetic have been adapted by taking into account majorizability considerations. One of such adaptations, the monotone functional interpretation of Ulrich Kohlenbach, has been at the center of a vigorous program in applied proof theory dubbed *proof mining*. We discuss some of the traditional and majorizability interpretations, including the recent bounded interpretations, and focus on the main *theoretical* techniques behind proof mining.

Contents

This work is based on four lectures on proof interpretations given in *Days in Logic'06* in Coimbra (Portugal), January 2006. Meanwhile, these lectures were published under the title *Proof Interpretations* in "Days in Logic'06 (Two Tutorials)," eds. R. Kahle & I. Oitavem, Textos de Matemática (Série B), Departamento de Matemática, Universidade de Coimbra, vol. 36. pp. 1-42 (2006). The lectures have few proofs, except for the ones that are somewhat new, not well-known or not easily available (other proofs are given suitable bibliographic pointers). The present article is a revision of the published tutorial, with some changes and corrections, omitting some issues, and inserting new material presented in our address to the *Logic Colloquium'07*. An introduction was also especially written for the present publication. The Coimbra lectures are rather hard to find (and came out with some annoying misprints). We hope that with the present publication our lectures become more widely available. We also want to thank the two anonymous referees for their suggestions of improvement. This work was partially supported by Fundação para a Ciência e Tecnologia (Financiamento Base 2008 - ISFL/1/209).

1. Introduction

Functional interpretations were introduced half a century ago by Kurt Gödel in [17]. Gödel's interpretation uses functionals of finite type and is an *exact* interpretation. It is exact in the sense that it provides precise witnesses of existential statements. Another example of an exact (functional) interpretation is Georg Kreisel's modified realizability [40], [41]. In the last fifteen years or so, there has been an interest in interpretations which are not exact, but only demand *bounds* for existential witnesses. These interpretations, when dealing with bounds for functionals of every type, are based on the majorizability notions of William Howard [20] and Marc Bezem [4]. This is the case with Ulrich Kohlenbach's monotone modified realizability [29] and monotone functional interpretation [27], the bounded modified realizability [11] of Fernando Ferreira and Ana Nunes, or the bounded functional interpretation [11] of Ferreira and Paulo Oliva. There are, to be sure, other interpretations which also incorporate majorizability notions to a certain extent: for instance, the seminal Diller-Nahm interpretation [6], Wolfgang Burr's interpretation of $\mathsf{KP}\omega$ [5] or the very recent interpretation of Jeremy Avigad and Henry Towsner [3] which is able to provide a proof theoretic analysis of ID_1. The reader should consult [2], [53] and [43] for more information on proof interpretations. Specifically for Gödel's interpretation, I also suggest the articles which appeared in a recent issue of *dialectica* [48] commemorating the 50th anniversary of Gödel's paper.

In these lectures, we give an overview of the various functional interpretations based on the notion of majorizability of Howard/Bezem. The starting points are the exact interpretations of Gödel and Kreisel. We have tried to organize the results around certain main theorems: soundness, extraction, conservation and characterization theorems. We hope that this organization makes it easier to appreciate the differences between the various interpretations, their advantages and limitations, and also the techniques involved in proving the theorems (even though, as observed in the introductory footnote, the lectures have few proofs). We also pay special attention to certain issues. For instance, we discuss the advantages of the monotone functional interpretation vis-à-vis Gödel's interpretation (Section 4.2), and discuss the role of extensionality in relation to *intensional* majorizability and the uniform boundedness principles (Section 5.6). The lectures finish with an extraction result for the fully extensional classical theory $\mathsf{E\text{-}PA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^{1,1}$ (the result entails that this theory has the provably total functions of Peano Arithmetic) via a *false* theory by applying the techniques developed in Part 5. It is not known whether there is a direct route. In order not to distract the reader, we give almost no references during the exposition of the material. Each of the four lectures closes with a section on suggested readings and historical notes where we try to give the proper references.

The monotone functional interpretation and associated theorems have been used to guide the extraction of effective data from given ordinary proofs in mathematics. In more picturesque words, it has been used to guide the "proof mining" of ordinary mathematical arguments. This paper does not attempt to even give a brief description of the applied work apart from saying that proof mining techniques (functional style) have been applied to approximation theory, fixed points of non-expansive mappings and, more recently, to ergodic theory. There are very good places to get into these applications and the recent book of Kohlenbach [35] is a very good place to start. Nevertheless, we would like to convey here a sense of the excitement of possible applications. Our example is taken from the research blog of Terence Tao. In a page entitled "Soft analysis, hard analysis, and the finite convergence principle" (see [51]), Tao speaks of an informal distinction between "soft" and "hard" principles in mathematics. Typically, "soft" principles are abstract, infinitary and lack computational content. They are extremely useful in modern mathematics and, given the usual training of a mathematician, they are simple to state and easy to remember and apply. An example of a "soft" principle is (in Tao's terminology) the *infinite convergence theorem*: every bounded monotone sequence of real numbers is convergent. On the other hand, "hard" principles are concrete, finitary and with computational content. They seem to be more difficult to grasp than the corresponding "soft" principles. In his blog, Tao sets himself the task of finding the "hard" counterpart of the infinite convergence theorem. Actually, Tao considers the following modification of the infinite convergence principle: every bounded nondecreasing sequence of reals is a Cauchy sequence. (The reader of like mind knows that this modified statement is weaker than the original one in terms of set existence: see, for instance, section 13.3 of [35].) Tao proceeds in an informal, tentative manner, and is eventually satisfied with a principle dubbed the *finite convergence principle*. Afterwards – in an impressive application – Tao formulates a variant of the infinite convergence theorem for Hilbert spaces and shows that Szemerédi's regularity lemma, a major combinatorial tool in graph theory, follows from the corresponding finite convergence principle for Hilbert spaces.

As pointed by Towsner in a comment of the blog, the finite convergence principle can be seen as an application of Gödel's *dialectica* interpretation. In a detailed treatment, Kohlenbach shows in his recent book [35] that the finite convergence principle is the no-counterexample interpretation (cf. Section 4.5) of the modified infinite convergence principle reinforced with a uniformity observation (which is given an explicit quantitative meaning in Kohlenbach's analysis). Note that, for the case at hand, the no-counterexample interpretation coincides with the *dialectica* interpretation (after a double-negation translation). In the lines below, we show that the finite convergence principle is essentially the bounded functional interpretation (for the classical case) of the modified infinite convergence principle:[1,2] in contrast with the *dialectica* interpretation, the bounded functional interpretation takes care of the uniformities automatically. Let us see why this is so. We state the modified infinite convergence theorem in a normalized form as follows:

$$\forall k \in \mathbb{N} \forall x \in [0,1]^{\mathbb{N}} \exists N \in \mathbb{N} \forall n \in \mathbb{N}\, |x_{N+n} - x_N| \leq \tfrac{1}{2^k},$$

under the assumption that $x_i \leq x_{i+1}$, for all $i \in \mathbb{N}$. Alternatively, one can drop the assumption and work instead with the sequence $\tilde{x}_n = \max_{i \leq n} x_i$. This is what we will do. Rather than go through the moves of applying the syntactic transformation of the bounded functional interpretation, we will put the above statement in the

[1]A similar case can also be made for the monotone functional interpretation.

[2]In [8], we describe a direct bounded functional interpretation for the classical case. In the present lectures, the (intuitionistic) bounded functional interpretation is applied after a double-negation translation. The proviso "essentially" is made because, for the bounded functional interpretation to work as desired, the convergence principles would have to be formulated with *intensional* majorizability signs (these reformulations are nevertheless equivalent to the original statements by extensionality reasons, as discussed in the sequel and in Section 5.6). These are technical matters that we sidestep in this introduction.

form $\forall\exists$ (with an appropriate matrix) using the so-called *characteristic principles* associated with the bounded functional interpretation (for the classical case). These are the principles $\mathsf{bBAC}^{\omega}_{\unlhd}$ and $\mathsf{MAJ}^{\omega}_{\unlhd}$ of Theorem 5.12. The first principle is a version of choice, a consequence of which is:

$$\text{(I):} \quad \forall i \in \mathbb{N} \exists j \in \mathbb{N}\, A_{\exists}(i,j) \to \tilde{\exists} f \in \mathbb{N}^{\mathbb{N}} \forall i \in \mathbb{N} \exists j \leq f(i)\, A_{\exists}(i,j),$$

where $A_{\exists}$ is an existential formula, and the tilde on the quantifier "$\tilde{\exists} f$" means that f is nondecreasing. A combination of the first and second principle yields the following "collection" principle (in our application, it reduces to a "compactness" property):

$$\text{(II):} \quad \forall w \in [0,1]^{\mathbb{N}} \exists i \in \mathbb{N} A_{\exists}(w,i) \to \exists l \in \mathbb{N} \forall w \in [0,1]^{\mathbb{N}} \exists i \leq l A_{\exists}(w,i),$$

where $A_{\exists}$ is an existential formula. The quantification "$\forall w \in [0,1]^{\mathbb{N}}$" is a *bounded* quantification under a suitable representation of the compact separable metric space $[0,1]^{\mathbb{N}}$ (its topology is the product topology of $\mathbb{N}$-copies of the closed unit interval, and a natural metric is forthcoming). One should consult [35] for an exposition concerning representations of Polish spaces. According to the characteristic principles, the bounded quantification mentioned above should have been *intensional*, but in our application we can work with the regular bounded quantification because we apply it to a matrix $A_{\exists}(w,i)$ which is *extensional* in w (these issues are discussed in Section 5.6).[3]

Assume the modified infinite convergence principle. Consider the negation of the formula $\exists N \in \mathbb{N} \forall n \in \mathbb{N}\, |\tilde{x}_{N+n} - \tilde{x}_N| \leq \frac{1}{2^k}$. Using (I), this negation is equivalent to $\tilde{\exists} F \in \mathbb{N}^{\mathbb{N}} \forall N \in \mathbb{N} \exists n \leq F(N)\, |\tilde{x}_{N+n} - \tilde{x}_N| > \frac{1}{2^k}$ (note that the relation $>$ between real numbers is existential). We get,

$$\forall k \in \mathbb{N} \tilde{\forall} F \in \mathbb{N}^{\mathbb{N}} \forall x \in [0,1]^{\mathbb{N}} \exists N \in \mathbb{N} \forall n \leq F(N)\, |\tilde{x}_{N+n} - \tilde{x}_N| \leq \tfrac{1}{2^k}.$$

It is clear that the matrix "$\forall n \leq F(N)\, |\tilde{x}_{N+n} - \tilde{x}_N| \leq \frac{1}{2^k}$" can be replaced by "$\forall n \leq F(N)\, |\tilde{x}_{N+n} - \tilde{x}_N| < \frac{1}{2^k}$." The latter matrix can be considered existential. Hence, we can apply (II) and conclude that

$$\forall k \in \mathbb{N} \tilde{\forall} F \in \mathbb{N}^{\mathbb{N}} \exists M \in \mathbb{N} \forall x \in [0,1]^{\mathbb{N}} \exists N \leq M \forall n \leq F(N)\, |\tilde{x}_{N+n} - \tilde{x}_N| < \tfrac{1}{2^k}.$$

Observe that only the values of x_i, for $i < M + F(M) + 1$, do matter. Hence the above can the restated as,

$$\begin{cases} \forall k \in \mathbb{N} \tilde{\forall} F \in \mathbb{N}^{\mathbb{N}} \exists M \in \mathbb{N} \forall x \in [0,1]^{M+F(M)+1} \\ \qquad \exists N \leq M \forall n \leq F(N)\, |\tilde{x}_{N+n} - \tilde{x}_N| < \frac{1}{2^k}. \end{cases}$$

This is a straightforward reformulation of Tao's finite convergence principle. As usual, this "hard" principle is more difficult to read than the "soft" monotone convergence principle. We quote a description given by Tao of this principle: the

[3] In more familiar mathematical terms, our application of (II) boils down to the fact that the compact space $[0,1]^{\mathbb{N}}$ is covered by the family of open sets $U_i = \{w \in [0,1]^{\mathbb{N}} : A_{\exists}(w,i)\}$, with $i \in \mathbb{N}$. Hence, it has a finite sub-cover.

finite convergence principle "asserts that any sufficiently long (but finite) bounded monotone sequence will experience arbitrarily high-quality amounts of *metastability* with a specified error tolerance $\frac{1}{2^k}$, in which the duration $F(N)$ of the metastability exceeds the time N of onset of the metastability by an arbitrary function F which is specified in advance." As Tao says, this is significantly more verbose than the "soft" formulation.

In a more recent blog [50] entitled "The correspondence principle and finitary ergodic theory," Tao speaks of a *correspondence principle* between qualitative (or "soft") results in infinite dynamical systems and quantitative (or "hard") results in finite dynamical systems. He illustrates such correspondence principle with eight examples, all of them mediated by compactness properties. It would be interesting to see how Tao's examples are relate to the majorizability interpretations: are they essentially instances of the bounded functional interpretation? Moreover, the majorizability interpretations are based on majorizability properties, not on compactness properties (consult sections 17.7 and 17.8 of [35] to appreciate the difference, as well as the groundwork with new base types in [33] and [15]). Are there examples of "correspondences" waiting to be discovered based on the majorizability interpretations?

I want to thank my students Patrícia Engrácia, Gilda Ferreira and Jaime Gaspar for discussions and observations concerning the majorizability interpretations. I would also like to thank the two anonymous referees for their helpful criticism and suggestions of improvement.

2. Basic theory and some models

2.1. **The language of finite-type arithmetic.** The finite types $\mathcal{T}$ are syntactic expressions defined inductively: 0 (the base type) is a finite type; if τ and σ are finite types then $\tau \to \sigma$ is a finite type. It is useful to have the following interpretation in mind: the base type 0 is the type constituted by the natural numbers, whereas $\tau \to \sigma$ is the type of (total) functions of objects of type τ to objects of type σ.

To make the reading easier, we often omit brackets and associate the arrows to the right. E.g., $0 \to 0 \to 0$ means $0 \to (0 \to 0)$. The *pure types* are defined inductively for each natural number: the pure type corresponding to the natural number 0 is the base type 0; the pure type $n+1$ is $n \to 0$.

The language of Heyting arithmetic in all finite types, denoted by $\mathcal{L}_0^\omega$, is a sorted language with a sort for each finite type. There is a denumerable set of variables x^σ, y^σ, z^σ, etc for each type σ. When convenient, we omit the type superscript. There are two kinds of constants:

(a) *Logical constants* or *combinators*. For each pair of types ρ, τ there is a combinator of type $\rho \to \tau \to \rho$ denoted by $\Pi_{\rho,\tau}$; for each triple of types δ, ρ, τ there is a combinator of type

$(\delta \to \rho \to \tau) \to (\delta \to \rho) \to (\delta \to \tau)$
denoted by $\Sigma_{\delta,\rho,\tau}$.

(b) *Arithmetical constants.* The constant 0 of type 0; the *successor* constant S of type 1; for each type ρ, a *recursor* constant of type
$0 \to \rho \to (\rho \to 0 \to \rho) \to \rho$
denoted by R_ρ.

Constants and variables of type ρ are terms of type ρ. If t is a term of type $\rho \to \tau$ and q is a term of type ρ then $App(t, q)$ is a term of type τ. These are all the terms there are. A term with no variables is a *closed term.* We usually write tq or $t(q)$ for $App(t, q)$. When writing tqr without brackets we associate to the *left* (note the difference with the previous convention): $(t(q))(r)$. We also write $t(q, r)$ instead of $(t(q))(r)$. In general, $t(q, r, \ldots, s)$ stands for $(\ldots((t(q))(r))\ldots)(s)$.

Atomic formulas are formulas of the form $t = q$ where t and q are terms of type 0. Note that, in the present setting, there is only one primitive equality symbol (infixing between terms of type 0). *Formulas* are obtained from atomic formulas by means of the usual propositional connectives $\wedge$, $\vee$, $\to$, $\bot$ (*falsum*) and universal and existential quantifiers $\forall x^\sigma$ and $\exists x^\sigma$ for each type σ. As usual, $\neg A$ abbreviates $A \to \bot$ and $A \leftrightarrow B$ abbreviates $(A \to B) \wedge (B \to A)$.

2.2. **Heyting arithmetic in all finite types.** The theory HA_0^ω is based on *intuitionistic logic.* It also has the following axioms for combinators and equality:

(a) *Axioms for combinators.* $A[\Pi(x, y)/w] \leftrightarrow A[x/w]$ and $A[\Sigma(x, y, z)/w] \leftrightarrow A[x(z, yz)/w]$, where A is an atomic formula with a distinguished variable w and $A[t/w]$ is obtained from A by replacing the occurences of w by t.

(b) *Equality axioms.* $x = x$ (reflexivity); $x = y \wedge A[x/w] \to A[y/w]$, where A is an atomic formula with a distinguished (type zero) variable w.

and the arithmetical axioms:

(c) *Successor axioms.* $Sx \neq 0$ and $Sx = Sy \to x = y$.

(d) *Axioms for recursors.* $A[R(0, y, z)/w] \leftrightarrow A[y/w]$ and $A[R(Sx, y, z)/w] \leftrightarrow A[z(R(x, y, z), x)/w]$, where A is an atomic formula with distinguished variable w.

(e) *Induction scheme.* $A(0) \wedge \forall x^0(A(x) \to A(Sx)) \to \forall x A(x)$, for each formula A of the language.

It can be shown that equality is symmetric and transitive and, moreover, the conditional $x = y \wedge A[x/w] \to A[y/w]$ also holds for every formula of the language (provided that there is no clash of variables). Similarly, the axioms for combinators and recursors extend to every formula A.

Equality is decidable in the following sense: $\mathsf{HA}_0^\omega \vdash \forall x^0(x = 0 \vee x \neq 0)$. This is easily proven by induction. Equality for higher types is defined inductively: $s =_{\rho\to\tau} t$ is $\forall x^\rho(sx =_\tau tx)$. Equality in higher types is not decidable anylonger.

Full extensionality is the scheme of axioms $\forall z \forall x, y(x = y \to zx = zy)$. It should be noted that we are *not* assuming full extensionality in the theory HA_0^ω. The theory obtained by its inclusion is denoted by $\mathsf{E\text{-}HA}_0^\omega$. In the sequel, we prefix a theory by E when adding full extensionality to it.

2.3. **Combinatorial completeness.** The combinators Π and Σ are instrumental in proving the following important property:

Theorem 2.1 (Combinatorial completeness)**.** *For each term $t[x^\rho]$ of type τ with a distinguished variable x of type ρ, we can construct a term q of type $\rho \to \tau$ whose variables are those of t except for x such that, for every term s of type ρ and atomic formula A with a distinguished variable w^τ,*

$$\mathsf{HA}_0^\omega \vdash A[t[s/x]/w] \leftrightarrow A[qs/w].$$

For instance, if t is of type 0, then $qs = t[s/x]$. The term q is usually denoted by $\lambda x.t[x]$ and the above equation can now be written $(\lambda x.t[x])s = t[s/x]$ (β-reduction). Of course, the proposition extends to all formulas A (provided that there is no clash of variables), i.e., we may substitute the term $t[s/x]$ by $(\lambda x.t[x])s$ in any formula.

Using the recursors, it is possible to associate to each description of a primitive recursive function a closed term that (with the proper understanding) satisfies the defining conditions of the description. Therefore, the theory HA_0^ω contains, in a natural sense, *primitive recursive arithmetic*. Actually, only the recursor R_0 is needed to define the primitive recursive functions. The presence of recursors of higher types has the effect of making possible the definition of functions beyond the primitive recursive functions (e.g., the Ackermann function).

We reserve the subscript "qf" for quantifier-free formulas:

Proposition 2.2. *For each quantifier-free formula $A_{\mathrm{qf}}(\underline{x})$ there is a closed term t of appropriate type such that $\mathsf{HA}_0^\omega \vdash A_{\mathrm{qf}}(\underline{x}) \leftrightarrow t\underline{x} = 0$.*

As a consequence, $\mathsf{HA}_0^\omega \vdash A_{\mathrm{qf}} \vee \neg A_{\mathrm{qf}}$. The theory obtained from HA_0^ω by adjoining the unrestricted law of excluded middle $A \vee \neg A$ is the classical theory PA_0^ω. If full extensionality is present, we have $\mathsf{E\text{-}PA}_0^\omega$.

2.4. **Main models.** 1. *The full set-theoretical model* S^ω. Let $\mathrm{S}_0 = \mathbb{N}$ and $\mathrm{S}_{\rho\to\tau} = (\mathrm{S}_\tau)^{\mathrm{S}_\rho}$, where $(\mathrm{S}_\tau)^{\mathrm{S}_\rho}$ is the set of *all* functions from S_ρ to S_τ. Let S^ω be $\langle \mathrm{S}_\sigma \rangle_{\sigma \in \mathcal{T}}$. With the proper understanding, it is clear that S^ω is a model of HA_0^ω. It is actually a model of $\mathsf{E\text{-}PA}_0^\omega$. The model S^ω is called the *standard* structure of finite-type arithmetic. When we call a sentence of $\mathcal{L}_0^\omega$ true or false, we always mean true or false with respect to the *standard* model.

2. *The hereditarily recursive operations* HRO^ω. For each type σ we define a subset HRO_σ of the natural numbers in the following way: $\mathrm{HRO}_0 = \mathbb{N}$ and

$$\mathrm{HRO}_{\rho\to\tau} = \{n \in \mathbb{N} : \forall k \in \mathrm{HRO}_\rho \, \exists m \in \mathrm{HRO}_\tau \, (\{n\}(k) \simeq m)\},$$

where $\{\cdot\}$ denotes the Kleene bracket of recursion theory. Let HRO^ω be $\langle \mathrm{HRO}_\sigma \rangle_{\sigma \in \mathcal{T}}$. If $n \in \mathrm{HRO}_{\rho\to\tau}$ and $k \in \mathrm{HRO}_\rho$ then $App^{\mathrm{HRO}^\omega}(n,k)$ is defined as $\{n\}(k)$. With a proper interpretation of the constants, HRO^ω is a model of PA_0^ω. By internalizing the above definitions within first-order Heyting arithmetic HA, it is possible to prove the following conservation result:

Proposition 2.3. *The theory* HA_0^ω *is conservative over* HA.

3. *The intensional continuous functionals* ICF^ω. In order to motivate some definitions below, let us briefly discuss continuous functionals of type 2. The set S_1 ($=\mathbb{N}^\mathbb{N}$) endowed with the product topology of the discrete space $\mathbb{N}$ with itself $\mathbb{N}$-times is known as the *Baire space*. It is easy to see that a functional $\Phi : \mathrm{S}_1 \mapsto \mathbb{N}$ is continuous at a point $\beta \in \mathrm{S}_1$ if, and only if,

$$\exists n \in \mathbb{N} \forall \gamma \in \mathrm{S}_1 (\overline{\gamma}(n) = \overline{\beta}(n) \to \Phi(\gamma) = \Phi(\beta)).$$

where $\overline{\gamma}(n)$ stands for the finite sequence $\langle \gamma(0), \ldots, \gamma(n-1)\rangle$. In other words, the value of Φ at a point β only depends on a finite initial segment of β. Therefore, Φ is determined by what happens in a countable set of finite sequences. This can be made explicit in the following way. An *associate* of a continuous functional Φ is an element $\alpha \in \mathrm{S}_1$ with the following two properties:

i. If $s = \langle s_0, \ldots, s_{n-1}\rangle$ is a finite sequence[4] of natural numbers, $\alpha(s) \neq 0$ and s is an initial segment of β, then $\Phi(\beta) = \alpha(s) - 1$.
ii. For all $\beta \in \mathrm{S}_1$, there is n such that $\alpha(\overline{\beta}(n)) \neq 0$.

Of course, such associates exist. It is clear that for every continuous functional Φ^2 with associate α, the following holds:

$$\Phi(\beta) = \alpha(\overline{\beta}(\mu k(\alpha(\overline{\beta}(k)) \neq 0))) - 1,$$

where μ is the minimization operator of recursion theory. Therefore, each continuous functional of type 2 is determined by a function of type 1. By means of this *type lowering* procedure, we obtain a structure for $\mathcal{L}_0^\omega$ in the following manner (for simplicity, we restrict ourselves to pure types). First, for $\alpha, \beta \in \mathrm{S}_1$ we define

$$\alpha(\beta) \simeq \alpha(\overline{\beta}(\mu k(\alpha(\overline{\beta}(k)) \neq 0))) - 1$$

Second, we let $\mathrm{ICF}_0 = \mathbb{N}$, $\mathrm{ICF}_1 = \mathrm{S}_1$ and, for the *remaining* pure cases:

$$\mathrm{ICF}_{\rho\to 0} = \{\alpha \in \mathrm{S}_1 : \forall \beta \in \mathrm{ICF}_\rho \exists k \in \mathbb{N}\, (\alpha(\beta) = k)\}$$

The above definition can be extended to all finite types. Let $\mathrm{ICF}^\omega = \langle \mathrm{ICF}_\sigma \rangle_{\sigma\in\mathcal{T}}$. Application between functionals is defined in the natural way: for $\alpha \in \mathrm{ICF}_{\rho\to 0}$ and $\beta \in \mathrm{ICF}_\rho$, with $\rho \neq 0$, then $App^{\mathrm{ICF}^\omega}(\alpha, \beta)$ is $\alpha(\beta)$. With a proper interpretation of the constants, ICF^ω is a model of PA_0^ω.

[4]We are identifying the finite sequence s with its numerical coding, as it is usually done in recursion theory.

This construction can be generalized in the following way. Take $U \subseteq \mathbb{N}^{\mathbb{N}}$ closed under Turing reducibility. Define $\mathrm{ICF}_0(U) = \mathbb{N}$, $\mathrm{ICF}_1(U) = U$ and, for $\sigma \neq 0, 1$, put $\mathrm{ICF}_\sigma(U)$ following the blueprint above (*mutatis mutandis*). Then $\mathrm{ICF}^\omega(U) := \langle \mathrm{ICF}_\sigma(U) \rangle_{\sigma \in \mathcal{T}}$ is a model of PA_0^ω. A particularly nice source of (counter-)examples is $\mathrm{ICF}^\omega(Rec)$, where Rec is the set of (total) recursive functions from $\mathbb{N}$ to $\mathbb{N}$. This is due to the fact that recursive sets do not necessarily separate disjoint r.e. sets (a result of Kleene).

4. *The extensional counterparts* HEO^ω *and* ECF^ω. The models HRO^ω and ICF^ω are called *intensional* because the scheme of extensionality fails to hold for them. For instance, take $e_1, e_2 \in \mathbb{N}$ two different indices for the constant zero function. Clearly, $e_1, e_2 \in \mathrm{HRO}_1$. Let $e \in \mathbb{N}$ be an index for the identity function. Of course, $e \in \mathrm{HRO}_2$. Even though $\{e\}(e_1) \neq \{e\}(e_2)$, one has $\forall n \in \mathrm{HRO}_0\,(\{e_1\}(n) = \{e_2\}(n))$. Therefore, the form of extensionality

$$\forall \Phi^2 \forall \alpha^1, \beta^1 (\forall k^0 (\alpha k = \beta k) \rightarrow \Phi\alpha = \Phi\beta),$$

already fails in HRO^ω. The above form of extensionality holds in ICF^ω, but it is easy to find a counter-example a type above, i.e. to:

$$\forall \mathfrak{G}^3 \forall \Phi^2, \Psi^2 (\forall \alpha^1 (\Phi\alpha = \Psi\alpha) \rightarrow \mathfrak{G}\Phi = \mathfrak{G}\Phi).$$

We now define the extensional counterparts of HRO^ω and ICF^ω. These are called, respectively, the *hereditarily effective operations* HEO^ω and the *(extensional) countinuous functionals* ECF^ω. The elements of non-zero type of HEO^ω and ECF^ω are functions (functionals). For HEO^ω we define by induction on the type σ both the functionals in HEO_σ and the *indices* of these functionals (these indices are natural numbers). We let HEO_0 be the set of natural numbers and declare that each natural number is an index of itself. We say that a natural number e is an index for a function (functional) $F : \mathrm{HEO}_\sigma \mapsto \mathrm{HEO}_\tau$ if for each index x of a function h of HEO_σ, $\{e\}(x)$ is defined and is an index for the function $F(h)$. We take $\mathrm{HEO}_{\rho\to\tau}$ as the set of functions $F : \mathrm{HEO}_\rho \mapsto \mathrm{HEO}_\tau$ which have an index. The structure HEO^ω is defined as $\langle \mathrm{HEO}_\sigma \rangle_{\sigma \in \mathcal{T}}$.

Regarding ECF^ω, we let ECF_0 be the set of natural numbers and, for each non-zero type σ, we define by induction both the functionals in ECF_σ and the *associates* of these functionals (these are elements of S_1). We put ECF_1 as S_1 and declare each $\alpha \in \mathrm{S}_1$ an associate of itself. We restrict to pure types in order to simplify. We say that $\alpha \in \mathrm{S}_1$ is an associate for a function $F : \mathrm{ECF}_\sigma \mapsto \mathbb{N}$, with σ a non-zero pure type, if for each associate β of a function h of ECF_σ, $\alpha(\beta)$ is defined and is the natural number $F(h)$. We take $\mathrm{ECF}_{\rho\to 0}$ as the set of functions $F : \mathrm{ECF}_\rho \mapsto \mathbb{N}$ which have an associate. This definition can be extended to all finite types. The structure ECF^ω is defined as $\langle \mathrm{ECF}_\sigma \rangle_{\sigma \in \mathcal{T}}$.

The structure ECF^ω is also called the structure of the *countable* functionals (Kleene's terminology). In a way similar to the intensional case, we can also start with any subset U of S_1 closed under Turing reducibility, and get the struture $\mathrm{ECF}^\omega(U)$.

5. *The strongly majorizable functionals* M^ω. Let $\mathrm{M}_0 = \mathbb{N}$ and let $\leq_0^*$ be the usual "less than or equal" relation between natural numbers. Given types ρ and τ, and given x and y elements of $\mathrm{M}_\tau^{\mathrm{M}_\rho}$, we say that $x \leq_{\rho\to\tau}^* y$ (and read "y strongly majorizes x") if

$$\forall u, v \in \mathrm{M}_\rho\,(u \leq_\rho^* v \to x(u) \leq_\tau^* y(v) \wedge y(u) \leq_\tau^* y(v))$$

Define $\mathrm{M}_{\rho\to\tau}$ as the set $\{x \in \mathrm{M}_\tau^{\mathrm{M}_\rho} : \exists y \in \mathrm{M}_\tau^{\mathrm{M}_\rho}(x \leq_{\rho\to\tau}^* y)\}$. Let M^ω be $\langle \mathrm{M}_\sigma \rangle_{\sigma\in\mathcal{T}}$. Bezem showed that M^ω forms a model of $\mathsf{E\text{-}PA}_0^\omega$ under function application and the natural interpretation of the constants. It is also important to observe that if $x, y \in \mathrm{M}_\tau^{\mathrm{M}_\rho}$ and $x \leq_{\rho\to\tau}^* y$, then $y \leq_{\rho\to\tau}^* y$. Therefore, not only is it the case that x is in $\mathrm{M}_{\rho\to\tau}$ but y is also there.

For $x, y \in \mathrm{S}_1$, $x \leq_1^* y$ iff $x \leq_1 y$ and y is non-decreasing. Note that $\leq_1^*$ is not reflexive. Given $\alpha : \mathbb{N} \mapsto \mathbb{N}$, let α^{M} be defined by $\alpha^{\mathrm{M}}(n) = \max_{k\leq n} \alpha(k)$. It is clear that $\alpha \leq_1^* \alpha^{\mathrm{M}}$. Therefore, $\mathrm{M}_1 = \mathrm{S}_1$. However, M_2 is *properly* contained in S_2. E.g., consider the functional $\Sigma \in \mathrm{S}_2$ defined as follows:

$$\Sigma(\alpha) = \begin{cases} n & \text{if } n \text{ is the least value such that } \alpha(n) \neq 0 \\ 0 & \text{if } \forall k\,(\alpha(k) = 0) \end{cases}$$

Suppose, in order to get a contradiction, that there is Ψ with $\Sigma \leq_2^* \Psi$. In particular, $\forall \alpha \in \mathrm{S}_1(\alpha \leq_1^* 1^1 \to \Sigma(\alpha) \leq \Psi(1^1))$. This is a contradiction: just consider the function α which takes the value 0 for numbers $n \leq \Psi(1^1)$ and is 1 afterwards. As a consequence, the following form of choice fails in M^ω: $\forall \alpha^1 \exists n^0 A(\alpha, n) \to \exists \Phi^2 \forall \alpha^1 A(\alpha, \Phi(\alpha))$. To see this, just take for A the formula:

$$(\alpha(n) \neq 0 \wedge \forall k < n(\alpha(k) = 0)) \vee (n = 0 \wedge \forall k(\alpha(k) = 0)).$$

The failure of choice in M^ω can be further improved by noticing that the *discontinuous* functional E defined thus:

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

is majorizable. Hence, we can replace the formula $\forall k(\alpha(k) = 0)$ by the quantifier-free formula $E(\alpha) = 1$ and, as a consequence, get the failure of the above form of choice with a quantifier-free matrix (although with the parameter E).

Continuous functionals have been playing an important role in discussions concerning finite-type functionals. Since the notion of majorizability (as opposed to continuity) plays the crucial role in some of the functional interpretations discussed in this paper, we believe that it is illuminating to make two observations regarding the relationship between continuity and majorizability. First, the continuous functionals of type 2 are in M_2. The proof of this fact uses a compactness argument. Given Φ a continuous functional of type 2, it makes sense to define the functional

$\Phi^{\mathrm{M}} \in \mathrm{S}_2$ according to the equation $\Phi^{\mathrm{M}}(\alpha) = \max_{\beta \leq_1 \alpha^{\mathrm{M}}} \Phi(\beta)$ because, for a given $\alpha \in \mathrm{S}_1$, the set $\{\beta \in \mathrm{S}_1 : \beta \leq_1 \alpha^{\mathrm{M}}\}$ is a *compact* subspace of the Baire space and, hence, $\Phi[\{\beta \in \mathrm{S}_1 : \beta \leq_1 \alpha^{\mathrm{M}}\}]$ is compact in the discrete topology of $\mathbb{N}$. In other words, it is finite and we can take its maximum. By construction, $\Phi \leq_2^* \Phi^{\mathrm{M}}$. Additionally, it is not difficult to show that Φ^{M} is continuous. Clearly, by the continuity of Φ one has

$$\mathrm{S}_1 = \cup_{s \in D}\{\beta \in \mathrm{S}_1 : \beta \text{ extends } s\},$$

where D is constituted by the finite sequences of natural numbers that determine values for Φ (i.e., such that elements of S_1 that extend an element of D have the same values according to Φ). Given $\alpha \in \mathrm{S}_1$, by compactness there is a finite $F_\alpha \subseteq D$ such that

$$\{\beta \in \mathrm{S}_1 : \beta \leq_1 \alpha^{\mathrm{M}}\} \subseteq \cup_{s \in F_\alpha}\{\beta \in \mathrm{S}_1 : \beta \text{ extends } s\}.$$

Let $n \in \mathbb{N}$ be the maximal length of the sequences in F_α. It can now be easily argued that if $\overline{\alpha}(n)$ is an initial segment of γ then $\Phi^{\mathrm{M}}(\gamma) = \Phi^{\mathrm{M}}(\alpha)$.

The second observation is that, nevertheless, there is a type 3 functional in ECF^ω which is not majorizable. Kreisel showed that there is a FAN functional $\mathfrak{F}^3$ in ECF_3 such that

$$\forall \Phi \in \mathrm{ECF}_2 \forall \alpha \leq_1 1 \forall \beta \leq_1 1\ (\alpha(\overline{\mathfrak{F}(\Phi)}) = \beta(\overline{\mathfrak{F}(\Phi)}) \to \Phi\alpha = \Phi\beta).$$

In other words, $\mathfrak{F}(\Phi)$ is a length witnessing the uniform continuity of Φ restricted to the Cantor space (i.e., to the compact subset of the Baire space constituted by the elements α such that $\alpha \leq_1 1^1$). However, the set $\{\mathfrak{F}(\Phi) : \Phi \in \mathrm{ECF}_2 \text{ and } \Phi \leq_2^* 1^2\}$ is clearly unbounded. Therefore, $\mathfrak{F}$ is not majorizable. (I thank Dag Normann for this observation.)

It will be important in the sequel to formalize the majorizability relation. We finish our discussion of the strongly majorizable functionals with this issue. Fix a suitable formula "$x \leq y$" of $\mathcal{L}_0^\omega$ saying that the natural number x is less than or equal to the natural number y. The (strong) majorizability formulas "$x \leq_\sigma^* y$" are defined inductively on the types according to the following clauses:

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

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

The following easy, but important, absoluteness property holds: when interpreted in M^ω the formulas "$x \leq_\sigma^* y$" coincide with the strong majorizability relations defined in the beginning of this example. It follows that M^ω is a model of the *majorizability axioms* MAJ^ω: $\forall x \exists y (x \leq^* y)$. By previous discussions, MAJ^ω already fails in the full set-theoretic model S^ω at type 2, and fails in the structure of (extensional) continuous functionals ECF^ω at type 3.

Lemma 2.4. *For each finite type σ, the theory HA_0^ω proves:*

(i) $x \leq_\sigma^* y \to y \leq_\sigma^* y$;

(ii) $x \leq^*_\sigma y \wedge y \leq^*_\sigma z \to x \leq^*_\sigma z$;
(iii) $x \leq_\sigma y \wedge y \leq^*_\sigma z \to x \leq^*_\sigma z$;

where the relation $\leq_\sigma$ *is the pointwise "less than or equal to" relation: it is* $\leq$ *for type 0, and* $x \leq_{\rho\to\tau} y$ *is defined recursively by* $\forall u^\rho\, (xu \leq_\tau yu)$.

The following theorem of Howard is a basic ingredient of the soundness proof of many interpretations discussed in this paper:

Theorem 2.5. *For each closed term* t *there is a closed term* q *such that* $\mathsf{HA}^\omega_0 \vdash t \leq^* q$.

6. *The term model.* We say that a term t *contracts* to a term q if one of the following clauses is satisfied:

(i) t is Πrs and q is r.
(ii) t is Σrst and q is $r(t, st)$.
(iii) t is $R0st$ and q is s.
(iv) t is $R(Sr)st$ and q is $t(Rrst, r)$.

A term t reduces to another term q in *one step* if q is obtained from t by contracting a single sub-term of t. A term is said to be in *normal form* if it does not admit reductions in one step. A *reduction sequence* is a sequence of terms $t_0, \ldots, t_n$ such that each term reduces in one step to the next. In this case, we say that t_0 *reduces* to t_n. If t_n cannot be reduced further, then t_n is in normal form and the sequence is called *terminating.*

Theorem 2.6 (Confluence and strong normalization)**.** *Every term reduces to a unique term in normal form. Moreover, every reduction sequence eventually terminates.*

The uniqueness result (confluence) is the Church-Rosser theorem for this reduction calculus. Normalization *strictu sensu* is the fact that every term has a terminating reduction sequence (*strong* normalization is the fact that *every* reduction sequence eventually terminates). Proofs of normalization are bound to use strong forms of induction (viz. induction on non-arithmetical predicates), because it is known that it (elementarily) implies the consistency of first-order Peano arithmetic. This fact can be proved with the aid of Gödel's *dialectica* interpretation (see Section 4.6).

A corollary of the normalization theorem is that every closed term r of type 0 reduces to a numeral $\overline{n}$ and, clearly, $\mathsf{HA}^\omega_0 \vdash r = \overline{n}$. Therefore, we can assign to each closed term t of type 1 a number theoretical function that maps each $k \in \mathbb{N}$ to the unique natural number n such that $t\overline{k}$ reduces to $\overline{n}$. The normalization process ensures that this function is recursive. As we will comment in Part 3, closed terms of type 1 give exactly the *provably total* Σ^0_1-functions of PA^ω_0 (and, actually, of PA).

Let CT_σ be the set of *closed terms* of type σ, and let CT^ω be $\langle \mathrm{CT}_\sigma \rangle_{\sigma \in \mathcal{T}}$. Given closed terms t, q of type zero, we say that $t =^{\mathrm{CT}^\omega} q$ if the normal forms of t and q are the same term. If $t \in \mathrm{CT}_{\rho \to \tau}$ and $q \in \mathrm{CT}_\rho$ then $App^{\mathrm{CT}^\omega}(t, q)$ is defined as the *term* $App(t, q)$. With these specifications and the interpretation of constants by themselves, CT^ω is a model of HA_0^ω.

A modification of the reduction calculus above has an interesting application. Let T be a closed term of type 2. The interpretation of T in the full set-theoretical model S^ω, denoted by T^{S^ω}, is a function from S_1 to $\mathbb{N}$. We claim that this function is continuous. In fact, we claim more: T^{S^ω} is a computable function (such functions are, of necessity, continuous since each computation only depends on a finite initial segment of the input, which is considered an oracle for effecting the computation). Consider a distinguished variable $\check{x}$ of type 1. Fix $\alpha \in \mathrm{S}_1$. We complement the previous normalization calculus with a new contraction rule:

$(\mathrm{v})_\alpha$ t is $\check{x}(\overline{n})$ and q is $\overline{\alpha(n)}$.

This α-normalization calculus enjoys the property of confluence and strong normalization. Therefore, the type 0 term $T\check{x}$ has a (unique) normal form which (it may easily be argued) must be a numeral $\overline{n}$. Due to the semantical soundness of the contraction procedure, it is clear that $T^{\mathrm{S}^\omega}(\alpha) = n$. It is also clear that the process of normalization yields an oracle computation (the oracle is only invoked to effect contractions of the form $(\mathrm{v})_\alpha$).

2.5. **Suggested reading and historical notes.** The structure of the hereditarily continuous functionals was independently discovered by Stephen Kleene in [22] and Kreisel in [40]. The notion of majorizability was introduced by Howard in [20]. The structure of the *strongly* majorizable functionals was defined by Bezem in [4]. The failure of choice in M^ω is due to Kohlenbach in [25]. The existence of the continuous FAN functional mentioned at the end of the discussion of the structure M^ω appeared in [41], and the proof that this functional is not majorizable is discussed in [23]. The strong normalization theorem for finite-type functionals (as well as the oracle modification) is due to William Tait in [49].

Anne Troelstra's book [55] is still very much rewarding for studying the topics of this section. Volume II of [54] and the recent [35] are alternatives. More specifically: HRO^ω and ICF^ω and their extensional counterparts are covered in [55] ([54] has less material). Our treatment of the extensional structures is slightly unusual since we work directly with set theoretic functionals instead of working with their indexes or associates endowed with a suitable notion of equality. Kohlenbach also studies also studies the extensional counterparts in [35], as well as the structure of the majorizable functionals in detail. Both [55] and [54] study the term model and the normalization theorems.

The treatment of equality in finite types in this paper comes from a suggestion in [18], elaborated by Troelstra in [52] (the subscript '0' in HA_0^ω comes from Troelstra given notation).

3. Modified and bounded realizability

3.1. **Modified realizability.** The method of realizability is reminiscent of the BHK (Brouwer-Heyting-Kolmogorov) interpretation of the intuitionistic connectives. We introduce a version of realizability, called *modified realizability*, due to Kreisel in the setting of finite-type arithmetic.

We present Kreisel's modified realizability in a slightly unfamilar way. Instead of saying what realizing tuples of functionals are, we associate to each formula of the language an existential formula. We need a preliminary definition:

Definition 3.1. *A formula of $\mathcal{L}_0^\omega$ is called $\exists$-free if it is built from atomic formulas by means of conjuntions, implications and universal quantifications.*

Do notice that $\exists$-free also means free of disjunctions (compare with some definitions in the sequel). We are ready to define the assignment of modified realizability:

Definition 3.2. *To each formula A of the language $\mathcal{L}_0^\omega$ we assign formulas $(A)^{\mathrm{mr}}$ and A_{mr} so that $(A)^{\mathrm{mr}}$ is of the form $\exists \underline{x} A_{\mathrm{mr}}(\underline{x})$ with $A_{\mathrm{mr}}(\underline{x})$ a $\exists$-free formula, according to the following clauses:*

1. *$(A)^{\mathrm{mr}}$ and A_{mr} are simply A, for atomic formulas A.*

If we have already interpretations for A and B given by $\exists \underline{x} A_{\mathrm{mr}}(\underline{x})$ and $\exists \underline{y} B_{\mathrm{mr}}(\underline{y})$ (respectively), then we define:

2. *$(A \wedge B)^{\mathrm{mr}}$ is $\exists \underline{x}, \underline{y}(A_{\mathrm{mr}}(\underline{x}) \wedge B_{\mathrm{mr}}(\underline{y}))$,*
3. *$(A \vee B)^{\mathrm{mr}}$ is $\exists n^0 \exists \underline{x}, \underline{y}\big((n = 0 \to A_{\mathrm{mr}}(\underline{x})) \wedge (n \neq 0 \to B_{\mathrm{mr}}(\underline{y}))\big)$,*
4. *$(A \to B)^{\mathrm{mr}}$ is $\exists \underline{f} \forall \underline{x}(A_{\mathrm{mr}}(\underline{x}) \to B_{\mathrm{mr}}(\underline{f}(\underline{x})))$,*
5. *$(\forall z A(z))^{\mathrm{mr}}$ is $\exists \underline{f} \forall z A_{\mathrm{mr}}(\underline{f}(z), z)$,*
6. *$(\exists z A(z))^{\mathrm{mr}}$ is $\exists z, \underline{x} A_{\mathrm{mr}}(\underline{x}, z)$.*

In the established literature, we say that $\underline{x}$ mr-*realizes* A instead of $A_{\mathrm{mr}}(\underline{x})$. Notice that the tuple $\underline{x}$ may be empty. The realizers of a disjunction include a flag n of type 0 that decides which way to fork. Similarly, the realizers of an existential quantifier include an existential witness. It is easy to check that the interpretation of negation $(\neg A)^{\mathrm{mr}}$ is $\forall \underline{x} \neg A_{\mathrm{mr}}(\underline{x})$. Notice that $(\neg A)^{\mathrm{mr}}$ is always an $\exists$-free formula and, therefore, demands an empty realizer. Realizability is unsuitable for extracting constructive information from negated formulas.

There are two important principles in connection with modified realizability:

I. *Axiom of Choice* AC^ω: $\forall \underline{x} \exists \underline{y} A(\underline{x}, \underline{y}) \to \exists \underline{f} \forall \underline{x} A(\underline{x}, \underline{f}\underline{x})$, where A is any formula.

II. *Independence of Premises* $\mathsf{IP}^{\omega}_{\exists\text{free}}$: $(A \to \exists \underline{z} B(\underline{z})) \to \exists \underline{z}(A \to B(\underline{z}))$, where A is ∃-free and B is an arbitrary formula.

The last principle is a classical, but not intuitionistic, law.

Theorem 3.3 (Soundness of modified realizability)**.** *Suppose that*

$$\mathsf{HA}^{\omega}_{0} + \mathsf{AC}^{\omega} + \mathsf{IP}^{\omega}_{\exists\text{free}} + \Delta \vdash A(\underline{z}),$$

where Δ *is a set of* ∃*-free sentences and* $A(\underline{z})$ *is an arbitrary formula (with the free variables as shown). Then there are closed terms* $\underline{t}$ *of appropriate types such that*

$$\mathsf{HA}^{\omega}_{0} + \Delta \vdash \forall \underline{z} A_{\text{mr}}(\underline{t}(\underline{z}), \underline{z}).$$

The proof is by induction on the length of formal derivations, and the terms are effectively constructed from the formal derivations. The principles AC^{ω} and $\mathsf{IP}^{\omega}_{\exists\text{free}}$ disappear because they are trivially realizable. It must be noted that the recursors are only needed to check the induction axioms, while the logical axioms only need terms built by the combinators. Full extensionality is not automatically interpretable but the next best thing happens: It is constituted by ∃-free sentences and, therefore, it is self-interpretable. Therefore, in the above theorem, we may replace in *both* places the theory HA^{ω}_{0} by $\mathsf{E}\text{-}\mathsf{HA}^{\omega}_{0}$.

3.2. **Extraction and all that (I).** The following is an important consequence of the soundness theorem:

Proposition 3.4 (Extraction and conservation, modified realizability)**.** *Suppose that*

$$\mathsf{HA}^{\omega}_{0} + \mathsf{AC}^{\omega} + \mathsf{IP}^{\omega}_{\exists\text{free}} + \Delta \vdash \forall x \exists y A(x, y),$$

where Δ *is a set of* ∃*-free sentences and* A *is a* ∃*-free formula with free variables among* x *and* y*. Then there is a closed term* t *of appropriate type such that*

$$\mathsf{HA}^{\omega}_{0} + \Delta \vdash \forall x A(x, tx).$$

Letting A be the sentence $0 = 1$ and Δ be empty, we conclude that the theory $\mathsf{HA}^{\omega}_{0} + \mathsf{AC}^{\omega} + \mathsf{IP}^{\omega}_{\exists\text{free}}$ is consistent *relative* to HA^{ω}_{0}. This is not, however, terribly interesting.

Theorem 3.5 (Characterization)**.** *For any formula* A*,*

$$\mathsf{HA}^{\omega}_{0} + \mathsf{AC}^{\omega} + \mathsf{IP}^{\omega}_{\exists\text{free}} \vdash A \leftrightarrow (A)^{\text{mr}}.$$

The two principles AC^{ω} and $\mathsf{IP}^{\omega}_{\exists\text{free}}$ are called the *characteristic principles* of modified realizability. The characterization theorem also ensures that we are not missing any principles besides AC^{ω} and $\mathsf{IP}^{\omega}_{\exists\text{free}}$ in the statement of the soundness theorem. To see this, suppose that we could state the soundness theorem with a further principle (sentence) P. Since P is a consequence of itself, from soundness it would follow that there are closed terms $\underline{t}$ such that $\mathsf{HA}^{\omega} \vdash \mathsf{P}_{\text{mr}}(\underline{t})$. *A fortiori*,

$\mathsf{HA}^\omega \vdash \exists \underline{x} \mathsf{P}_{\mathrm{mr}}(\underline{x})$, i.e., $\mathsf{HA}^\omega \vdash (\mathsf{P})^{\mathrm{mr}}$. By the characterization theorem, we get $\mathsf{HA}^\omega + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash \mathsf{P}$. In conclusion, P is superfluous.

Corollary 3.6. *The following two properties hold:*

(a) *(Disjunction Property) Suppose that* $\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash A \vee B$*, for* sentences A *and* B*. Then, either* $\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash A$ *or* $\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash B$*.*

(b) *(Existence Property) Suppose that* $\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash \exists x A(x)$*, for the* sentence $\exists x A(x)$*. Then there is a closed term* t *such that* $\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash A(t)$*.*

The above corollary is consequence of the soundness and characterization theorems. For instance, suppose that $A \vee B$ is provable in $\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}}$. By the soundness theorem, there are closed terms s^0 and $\underline{t}, \underline{q}$ such that,

$$\mathsf{HA}^\omega_0 \vdash (s = 0 \to A_{\mathrm{mr}}(\underline{t})) \wedge (s \neq 0 \to B_{\mathrm{mr}}(\underline{q})).$$

Since s is a closed term, there is a numeral $\overline{n}$ such that $\mathsf{HA}^\omega_0 \vdash s = \overline{n}$. Suppose $n = 0$ (the other case is similar). Then $\mathsf{HA}^\omega_0 \vdash A_{\mathrm{mr}}(\underline{t})$ and, therefore, $\mathsf{HA}^\omega_0 \vdash (A)^{\mathrm{mr}}$. By the characterization theorem, we get the desired conclusion.

It is important to notice that the above argument works because, in the presence of AC^ω and $\mathsf{IP}^\omega_{\exists\mathrm{free}}$, it is possible to *come back* from the formula $(A)^{\mathrm{mr}}$ to the original formula A. The existence property has the following natural generalization:

Proposition 3.7. *Let* A *be an* arbitrary *formula with free variables among* x *and* y*. Suppose that*

$$\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash \forall x \exists y A(x, y).$$

Then there is a closed term t *of appropriate type such that*

$$\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash \forall x A(x, tx).$$

This is not a conservation result anymore. However, it is still a *sound* extraction result since the conclusion $\forall x A(x, tx)$ is *true.*

Theorem 3.8 (Fan rule)**.** *Let* A *be an arbitrary formula containing only the variables* x^1 *and* n^0*. Then the following rule holds: If*

$$\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash \forall x \leq_1 1 \exists n^0 \, A(x, n)$$

then there is a natural number m *such that*

$$\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash \forall x \leq_1 1 \exists n \leq \overline{m} \, A(x, n).$$

The proof of this result uses Howard's majorizability relation. Assume that $\forall x \leq_1 1 \exists n^0 \, A(x, n)$ is provable in $\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}}$. We get,

$$\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_{\exists\mathrm{free}} \vdash \forall x (x \leq_1 1 \to \exists n \exists \underline{z} A_{\mathrm{mr}}(\underline{z}, x, n)).$$

By $\mathsf{IP}^\omega_{\exists\mathrm{free}}$, $\forall x \exists n \exists \underline{z} (x \leq_1 1 \to A_{\mathrm{mr}}(\underline{z}, x, n))$ is provable. As a consequence of the soundness theorem, there is a closed term t of type 2 such that

$$\mathsf{HA}_0^\omega \vdash \forall x(x \leq_1 1 \to \exists \underline{z} A_{\mathrm{mr}}(\underline{z}, x, tx)).$$

By Howard's majorizability result, take a closed term q such that $t \leq_2^* q$. It is clear that

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_{\exists\mathrm{free}}^\omega \vdash \forall x \leq_1 1 \exists n \leq q(1^1)\, A(x, n).$$

3.3. **Digression on intuitionistic extraction.** Modified realizability does not quite achieve the disjunction and existence property for HA_0^ω. For instance, if $\mathsf{HA}_0^\omega \vdash A \vee B$ then we can only guarantee that either A or B is provable in the *stronger* theory $\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_{\exists\mathrm{free}}^\omega$. One would want instead that A or B were already provable in the original HA_0^ω. This is in fact the case, and the proof uses a variant of modified realizability: *modified realizability with truth.* The clauses of this variant of realizability are the same of modified realizability except that the conditional has an extra clause:

$$(A \to B)^{\mathrm{mrt}} \textit{ is } \exists \underline{f} \forall \underline{x}(A_{\mathrm{mrt}}(\underline{x}) \to B_{\mathrm{mrt}}(\underline{f}(\underline{x}))) \wedge (A \to B)$$

Theorem 3.9 (Soundness of modified realizability with truth)**.** *Suppose that*

$$\mathsf{HA}_0^\omega \vdash A(\underline{z}),$$

where $A(\underline{z})$ is an arbitrary formula (with the free variables as shown). Then there are closed terms $\underline{t}$ of appropriate types such that

$$\mathsf{HA}_0^\omega \vdash \forall \underline{z} A_{\mathrm{mrt}}(\underline{t}(\underline{z}), \underline{z}).$$

For the purpose at hand, we crucially have:

Lemma 3.10. *For every formula A, $\mathsf{HA}_0^\omega \vdash (A)^{\mathrm{mrt}} \to A$.*

Now it is clear how to prove the disjunction and existence properties for plain HA_0^ω.

3.4. **Bounded modified realizability.** The new bounded interpretations rely heavily on the Howard-Bezem majorizability notions. In view of this fact, it is convenient to work with an extension of the language $\mathcal{L}_0^\omega$ (with exactly the same terms). Firstly, we extend the language $\mathcal{L}_\leq^\omega$ with a *primitive* binary relation symbol $\leq$ that infixes between terms of type 0. There are now new atomic formulas, and the syntactic notions extend in the natural way. Actually, the language that we get is an extension *by definitions* of $\mathcal{L}_0^\omega$ because $x \leq y$ may be defined by a natural quantifier-free formula. In this setting, we use the primitive binary symbol to define the majorizability formulas. Secondly, it is convenient to introduce the primitive syntactical device of *bounded quantifications*, i.e., quantifications of the form $\forall x \leq^* t$ and $\exists x \leq^* t$, for terms t not containing the variable x. *Bounded formulas* are formulas in which every quantifier is bounded.

The theory $\mathsf{HA}_\leq^\omega$ is HA_0^ω together with the (universal) defining axiom of $\leq$ and the following schemes:

$\mathsf{B}_\forall$: $\forall x \leq^* t A(x) \leftrightarrow \forall x (x \leq^* t \rightarrow A(x))$,
$\mathsf{B}_\exists$: $\exists x \leq^* t A(x) \leftrightarrow \exists x (x \leq^* t \wedge A(x))$.

It is clear that $\mathsf{HA}^\omega_\leq$ is a conservative extension of HA^ω_0. We say that a functional f is monotone if $f \leq^* f$. In the sequel, we often quantify over monotone functionals. We abbreviate the quantifications $\forall x (x \leq^* x \rightarrow A(x))$ and $\exists x (x \leq^* x \wedge A(x))$ by $\tilde{\forall} x A(x)$ and $\tilde{\exists} x A(x)$, respectively.

Definition 3.11. *A formula of $\mathcal{L}^\omega_\leq$ is called* $\tilde{\exists}$-free *if it is built from atomic formulas by means of conjunctions, disjunctions, implications, bounded quantifications and monotone universal quantifications, i.e., quantifications of the form $\tilde{\forall} a$.*

This notion resembles the notion of $\exists$-free formula, but notice that *disjunctions* are allowed.

Definition 3.12. *To each formula A of the language $\mathcal{L}^\omega_\leq$ we assign formulas $(A)^{\mathrm{br}}$ and A_{br} so that $(A)^{\mathrm{br}}$ is of the form $\tilde{\exists} \underline{b} A_{\mathrm{br}}(\underline{b})$, with $A_{\mathrm{br}}(\underline{b})$ a $\tilde{\exists}$-free formula, according to the following clauses:*

1. *$(A)^{\mathrm{br}}$ and $(A)_{\mathrm{br}}$ are simply A, for atomic formulas A.*

If we have already interpretations for A and B given by $\tilde{\exists} \underline{b} A_{\mathrm{br}}(\underline{b})$ and $\tilde{\exists} \underline{d} B_{\mathrm{br}}(\underline{d})$ (respectively) then, we define

2. *$(A \wedge B)^{\mathrm{br}}$ is $\tilde{\exists} \underline{b}, \underline{d} (A_{\mathrm{br}}(\underline{b}) \wedge B_{\mathrm{br}}(\underline{d}))$,*
3. *$(A \vee B)^{\mathrm{br}}$ is $\tilde{\exists} \underline{b}, \underline{d} (A_{\mathrm{br}}(\underline{b}) \vee B_{\mathrm{br}}(\underline{d}))$,*
4. *$(A \rightarrow B)^{\mathrm{br}}$ is $\tilde{\exists} \underline{f} \tilde{\forall} \underline{b} (A_{\mathrm{br}}(\underline{b}) \rightarrow B_{\mathrm{br}}(\underline{f}(\underline{b})))$.*

For bounded quantifiers we have:

5. *$(\forall x \leq^* t\, A(x))^{\mathrm{br}}$ is $\tilde{\exists} \underline{b} \forall x \leq^* t\, A_{\mathrm{br}}(\underline{b}, x)$,*
6. *$(\exists x \leq^* t\, A(x))^{\mathrm{br}}$ is $\tilde{\exists} \underline{b} \exists x \leq^* t\, A_{\mathrm{br}}(\underline{b}, x)$.*

And for unbounded quantifiers we define

7. *$(\forall x A(x))^{\mathrm{br}}$ is $\tilde{\exists} \underline{f} \tilde{\forall} a \forall x \leq^* a\, A_{\mathrm{br}}(\underline{f}(a), x)$.*
8. *$(\exists x A(x))^{\mathrm{br}}$ is $\tilde{\exists} a, \underline{b} \exists x \leq^* a\, A_{\mathrm{br}}(\underline{b}, x)$.*

Notice that the realizers of a disjunction do not include a flag deciding which way to fork, and that only a *bound* for the existential witness is included in the realizers of an existential statement. As usual, negation is a particular case of the implication: $(\neg A)^{\mathrm{br}}$ is $\tilde{\forall} \underline{b} \neg A_{\mathrm{br}}(\underline{b})$.

Three principles are important in connection with the above assignment:

I. *Bounded Choice* bAC^ω:

$$\forall \underline{x} \exists \underline{y} A(\underline{x}, \underline{y}) \rightarrow \tilde{\exists} \underline{f} \tilde{\forall} \underline{b} \forall \underline{x} \leq^* \underline{b} \exists \underline{y} \leq^* \underline{f}\underline{b}\, A(\underline{x}, \underline{y}),$$

where A is an arbitrary formula.

II. *Bounded Independence of Premises* $\mathsf{bIP}^\omega_{\tilde{\exists}\mathrm{free}}$:

$$(A \to \exists \underline{y} B(\underline{y})) \to \tilde{\exists} \underline{b}(A \to \exists \underline{y} \leq^* \underline{b}\, B(\underline{y})),$$

where A is a $\tilde{\exists}$-free formula and B is an arbitrary formula.

III. *Majorizability Axioms* MAJ^ω:

$$\forall x \exists y (x \leq^* y).$$

It must be remarked that each principle above is *false* in the full set theoretical model S^ω. The majorizability axioms MAJ^ω (as well as $\mathsf{bIP}^\omega_{\tilde{\exists}\mathrm{free}}$) are, of course, true in the structure M^ω of the majorizable functionals. However, as we saw in the first part, the bounded choice principle already fails in M^ω for x of type 1 and y of type 0.

Proposition 3.13. *The theory* $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\mathrm{free}}$ *proves the* Collection Principle bC^ω*:*

$$\forall \underline{z} \leq^* \underline{c}\, \exists \underline{y} A(\underline{y}, \underline{z}) \to \tilde{\exists} \underline{b} \forall \underline{z} \leq^* \underline{c} \exists \underline{y} \leq^* \underline{b} A(\underline{y}, \underline{z}),$$

where A is an arbitrary formula and $\underline{c}$ is a tuple of monotone functionals.

Proof. Suppose that $\forall \underline{z} \leq^* \underline{c}\, \exists \underline{y} A(\underline{y}, \underline{z})$. By $\mathsf{bIP}^\omega_{\tilde{\exists}\mathrm{free}}$, we get

$$\forall \underline{z} \tilde{\exists} \underline{b}(\underline{z} \leq^* \underline{c} \to \exists \underline{y} \leq^* \underline{b} A(\underline{y}, \underline{z})).$$

By bAC^ω, we may conclude that $\tilde{\exists} \underline{f} \forall \underline{z} \leq^* \underline{c} \tilde{\exists} \underline{b} \leq^* \underline{f}\underline{c} \exists \underline{y} \leq^* \underline{b}\, A(\underline{x}, \underline{y})$. The desired conclusion follows from the transitivity of the majorizability relation. □

The case where the types of z and y are 0 extends the familiar *collection principle* of arithmetic. In the context of intuitionistic analysis, Brouwer's FAN theorem is the case where z is of type 1 and y is of type 0. The formulation for the Cantor space is:

$$\forall z \leq_1 1 \exists n^0 A(n, z) \to \exists m^0 \forall z \leq_1 1 \exists n \leq m A(n, z),$$

for arbitrary formulas A. The above formulation of the FAN theorem must be distinguished from the following, which also appears in the literature:

$$\forall z \leq_1 1 \exists n^0 A(n, z) \to \exists k^0 \forall z \leq_1 1 \exists n^0 \forall w \leq_1 1 (\overline{w}(k) = \overline{z}(k) \to A(n, w)),$$

for arbitrary formulas A. The latter formulation explicitly includes a *continuity principle*, whereas the former does not. It is important to keep in mind that bounded interpretations concern majorizability notions, not continuity notions.

Theorem 3.14 (Soundness of the bounded modified realizability)**.** *Suppose that*

$$\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\mathrm{free}} + \mathsf{MAJ}^\omega + \Delta \vdash A(\underline{z}),$$

where Δ is a set of $\tilde{\exists}$-free sentences and $A(\underline{z})$ is an arbitrary formula (with the free variables as shown). Then there are closed monotone terms $\underline{t}$ of appropriate types such that

$$\mathsf{HA}^\omega_{\trianglelefteq} + \Delta \vdash \tilde{\forall} \underline{a} \forall \underline{z} \leq^* \underline{a}\, A_{\mathrm{br}}(\underline{t}(\underline{a}), \underline{z}).$$

Above, and in similar situations in the sequel, by a closed *monotone* term t we mean a closed term such that the monotonicity condition ($t \leq^* t$, in this case) is provable in the base theory ($\mathsf{HA}^\omega_\leq$, in this case).

3.5. **Extraction and all that (II).** The following is a consequence of the soundness theorem:

Proposition 3.15 (Extraction and conservation, bounded modified realizability)**.** *Suppose that*

$$\mathsf{HA}^\omega_\leq + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega + \Delta \vdash \forall x \exists y A(x, y),$$

where Δ is a set of $\tilde{\exists}$-free sentences and A is a $\tilde{\exists}$-free formula with free variables among x and y. Then there is a closed monotone term t of appropriate type such that

$$\mathsf{HA}^\omega_\leq + \Delta \vdash \tilde{\forall} a \forall x \leq^* a \exists y \leq^* ta\, A(x, y).$$

Strictly speaking, we do not have a conservation result. However, if the type of x is 0 or 1 then we do have such a result:

Corollary 3.16. *Suppose that*

$$\mathsf{HA}^\omega_\leq + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega + \Delta \vdash \forall x^{0/1} \exists y A(x, y),$$

where Δ is a set of $\tilde{\exists}$-free sentences, A a $\tilde{\exists}$-free formula with free variables among x and y. Then there is a closed monotone term t of appropriate type such that

$$\mathsf{HA}^\omega_\leq + \Delta \vdash \forall x \exists y \leq^* tx\, A(x, y).$$

In particular, this corollary yields a relative consistency result of the theory $\mathsf{HA}^\omega_\leq + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega$ over $\mathsf{HA}^\omega_\leq$. This is interesting in the present setting because the former theory is *classically inconsistent*: It refutes the classically true Markov's principle. To see this, suppose that

$$\forall x^1 \big(\neg\neg \exists n^0 (xn = 0) \to \exists n^0 (xn = 0)\big).$$

By intuitionistic logic and $\mathsf{bIP}^\omega_{\tilde{\exists}\text{free}}$, $\forall x^1 \exists n^0 (\neg \forall k^0 (xk \neq 0) \to \exists i \leq n (xi = 0))$. Now, by the collection principle bC^ω, there is a natural number m^0 such that $\forall x \leq_1 1 (\neg \forall k^0 (xk \neq 0) \to \exists n \leq m\, (xn = 0))$. This is a contradiction (just consider the number-theoretic primitive recursive function that takes the value 1 for values less than $m + 1$ and is 0 afterwards).

The reader should take notice that the world of $\mathsf{HA}^\omega_\leq + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega$ is a world with some principles related to Brouwerian intuitionism and, in the terminology of intuitionism, has *strong counterexamples* to classical logic.

We can also prove a characterization theorem for bounded modified realizability:

Theorem 3.17 (Characterization)**.** *For any formula A,*

$$\mathsf{HA}^\omega_\leq + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega \vdash A \leftrightarrow (A)^{\text{br}}.$$

Of course, bAC^ω, $\mathsf{bIP}^\omega_{\tilde{\exists}\text{free}}$ and MAJ^ω are called the *characteristic* principles of the bounded modified realizability. As noticed, in contrast with the traditional case, these charateristic principles are not true. If we prove $\forall x^1 \exists y A(x,y)$ in the extended theory $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega$ for an *arbitrary* A then, in complete analogy with the traditional case, we can find a closed monotone term t such that $\forall x^1 \exists y \leq^* tx\, A(x,y)$ is provable in $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega$. However, the extended theory is not a *true* theory and, therefore, we may not conclude that the statement $\forall x^1 \exists y \leq^* tx\, A(x,y)$ is true. In other words, the extraction of the term t is *not sound* when the matrix A is arbitrary. However, as we saw above, the extraction is indeed sound if $A(x,y)$ is an $\tilde{\exists}$-free formula (in particular, if $A(x,y)$ is quantifier-free).

3.6. **Benign principles.** In the context of bounded modified realizability, if we are interested in extracting sound bounding information from proofs, we must restrict ourselves to conclusions of the form $\forall\exists$ whose matrix is $\tilde{\exists}$-free. If such statements are proved in the theory $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega$, then it is possible to extract truthful bounding information. Of course, we may use in our proofs principles that *follow* from the above theory. This is the case, e.g., with Kohlenbach's *uniform boundedness principles* UB_ρ, a combination of the FAN theorem (extended to higher types) with choice:

$$\forall k^0 \forall x \leq yk\, \exists z^0 A(x,y,k,z) \to \exists \chi^1 \forall k^0 \forall x \leq yk\, \exists z \leq \chi k\, A(x,y,k,z),$$

for arbitrary A and y of type $0 \to \rho$.

There are, however, principles that do not follow from the extended theory but whose use as premises does yield sound bounding information (since that information is checked in a *true* theory). We call these principles *benign*. Note that benign principles may be *false* (because they could follow from suitable true principles with the aid of the false characteristic principles). We present a list of eight benign principles:

1. The axioms of extensionality $\forall z^{\sigma\to\delta} \forall x^\sigma, y^\sigma (x =_\sigma y \to zx =_\delta zy)$, for $\sigma = 0, 1$ or 2 and δ arbitrary, are benign.

2. The classical, but not intuitionistic, truth

$$\forall x \forall y (A(x) \vee B(y)) \to \forall x A(x) \vee \forall y B(y),$$

where A and B are $\tilde{\exists}$-free formulas, is benign. When the types of x and y are 0 and A and B are quantifier-free formulas, we have the *lesser limited principle of omniscience* LLPO (this is Errett Bishop's terminology).

3. The law of excluded middle $A \vee \neg A$, for $\tilde{\exists}$-free formulas A, is benign. This form of excluded middle includes $\Pi^0_1\text{−}\mathsf{LEM}$, i.e., $\forall n^0 A(n) \vee \neg \forall n^0 A(n)$ for A a first-order bounded formula. In view of the fact that $\mathsf{HA}^\omega_0 + \mathsf{bAC}^\omega + \mathsf{bIP}^\omega_{\tilde{\exists}\text{free}} + \mathsf{MAJ}^\omega$

refutes Markov's principle, we are drawn to the conclusion that $\Pi^0_1-\mathsf{LEM}$ does not prove Markov's principle.

4. The choice principle $\forall x^{0/1}\exists y A(x,y) \to \exists f \forall x A(x,fx)$, where A is an arbitrary formula and y is of any type, is benign.

5. The following is a benign version of choice with no restrictions on the types: $\forall x \leq^* a\, \exists y A(x,y) \to \exists f \forall x \leq^* a\, A(x,fx)$, for arbitrary formulas A.

6. It is well-known that the FAN theorem is *true* (and intuitionistically acceptable) for quantifier-free matrices with parameters of type 0 or 1 *only*. A modification of its contrapositive, which is equivalent to *weak König's lemma* WKL, is rejected intuitionistically. Nevertheless, this modified contrapositive, namely $\forall n^0 \exists x \leq_1 1 \forall k \leq n\, A_{\mathrm{qf}}(x,k) \to \exists x \leq_1 1 \forall k\, A_{\mathrm{qf}}(x,k)$, is a benign principle. The familiar formulation of weak König's lemma states that every infinite subtree of the full binary tree has an infinite path. This principle is non-constructive in the following precise sense: There are infinite recursive subtrees of the full binary tree that have no recursive infinite path (this is a reformulation of Kleene's result that there are recursively inseparable r.e. sets).

7. The form of comprehension $\exists \Phi \forall y (\Phi y = 0 \leftrightarrow A(y))$, where A a $\tilde{\exists}$-free formula and y may be of any type, is benign. Comprehension for negated formulas is also a benign principle: $\exists \Phi \forall y(\Phi y = 0 \leftrightarrow \neg A(y))$, for arbitrary A.

8. Kohlenbach considered in [29] the principles F_ρ, a simplification of which are:

$$\forall \Phi^{\rho \to 0} \forall y^\rho \exists y_0 \leq_\rho y \forall z \leq_\rho y\, (\Phi(z) \leq \Phi(y_0)).$$

These principles are false for $\rho \neq 0$. They are, nevertheless, benign.

The proof that the above principles are benign relies on a careful study of the formulas which imply, or are implied by, their own bounded realizations.

For some years now, Kohlenbach and his co-workers have been showing the *practical* use of Proof Theory in obtaining numerical bounds from *classical* proofs of analysis. Kohlenbach's methods are *not* based on realizability because realizability notions (including bounded realizability) are *not* taylored for the analysis of *classical* proofs. In effect, even though a classical proof may be translated into an intuitionistic proof via (e.g.) the Gödel-Gentzen negative translation, the translation destroys existential statements – replacing them by negated universal statements – with the consequence that realizers yield no computational information. Of course, this shortcoming is related with the fact that Markov's principle is *not* benign. That notwithstanding, bounded modified realizability (and Kohlenbach's monotone modified realizability) supports many classical principles that go beyond intuitionistic logic.

3.7. **Suggested reading and historical notes.** The notion of (numerical) realizability was introduced by Stephen Kleene in [21]. Modified realizability is due to Kreisel in [40] and [41]. Proofs of the theorems of Section 3.1 can be found in [55]. An alternative is the very recent [35] (which has, nevertheless, circulated in preliminary form as a manuscript for quite some years). A survey on realizability until the mid-nineties can be found in [53] (where one can find the story of the truth variants of realizability). The FAN rule at the end of Section 3.2 appeared in [35] but is already essentially treated in [25]. Bounded modified realizability was introduced by Ferreira and Nunes in [11]. This paper includes full proofs of the results concerning this form of realizability, including the discussion of the benign principles. For the sake of space, we did not discuss the monotone version of realizability introduced by Kohlenbach in [29]. The result that $\Pi^0_1-\mathsf{LEM}$ does not prove Markov's principle first appeared in [1] using this version. Monotone modified realizability is also explained in [35].

4. The *dialectica* interpretation

4.1. Gödel's *dialectica* interpretation.

In 1958, Gödel published an article in an issue of the journal *dialectica* dedicated to the seventieth anniversary of Paul Bernays. The article presents an interpretation of first-order Heyting arithmetic into a quantifier-free theory with finite-type functionals (Gödel's theory T). This theory is, essentially, the quantifier-free part of HA^ω_0. Gödel's *dialectica* interpretation appeared as a contribution to an extended Hilbert's program by means of the notion of computable functional of finite type. In the sequel, we present Gödel's interpretation extended to finite-type arithmetic HA^ω_0.

Definition 4.1. *To each formula A of the language $\mathcal{L}^\omega_0$ we assign formulas $(A)^{\mathrm{D}}$ and A_{D} so that $(A)^{\mathrm{D}}$ is of the form $\exists \underline{x} \forall \underline{y} A_{\mathrm{D}}(\underline{x}, \underline{y})$ with $A_{\mathrm{D}}(\underline{x}, \underline{y})$ a quantifier-free formula, according to the following clauses:*

1. *$(A)^{\mathrm{D}}$ and A_{D} are simply A, for atomic formulas A.*

If we have already interpretations for A and B given by $\exists \underline{x} \forall \underline{y} A_{\mathrm{D}}(\underline{x}, \underline{y})$ and $\exists \underline{z} \forall \underline{w} B_{\mathrm{D}}(\underline{z}, \underline{w})$ (respectively) then we define:

2. *$(A \wedge B)^{\mathrm{D}}$ is $\exists \underline{x}, \underline{z} \forall \underline{y}, \underline{w}(A_{\mathrm{D}}(\underline{x}, \underline{y}) \wedge B_{\mathrm{D}}(\underline{z}, \underline{w}))$,*
3. *$(A \vee B)^{\mathrm{D}}$ is $\exists n^0, \underline{x}, \underline{z} \forall \underline{y}, \underline{w}((n = 0 \to A_{\mathrm{D}}(\underline{x}, \underline{y})) \wedge (n \neq 0 \to B_{\mathrm{D}}(\underline{z}, \underline{w})))$,*
4. *$(A \to B)^{\mathrm{D}}$ is $\exists \underline{f}, \underline{g} \forall \underline{x}, \underline{w}(A_{\mathrm{D}}(\underline{x}, \underline{g}(\underline{x}, \underline{w})) \to B_{\mathrm{D}}(\underline{f}(\underline{x}), \underline{w}))$.*
5. *$(\forall z A(z))^{\mathrm{D}}$ is $\exists \underline{f} \forall z \forall \underline{y} A_{\mathrm{D}}(\underline{f}(z), \underline{y}, z)$.*
6. *$(\exists z A(z))^{\mathrm{D}}$ is $\exists z, \underline{x} \forall \underline{y} A_{\mathrm{D}}(\underline{x}, \underline{y}, z)$.*

The definition of implication is the hardest to understand. Gödel motivates it as follows. First consider $(A)^{\mathrm{D}} \to (B)^{\mathrm{D}}$, that is, the implication $\exists \underline{x} \forall \underline{y} A_{\mathrm{D}}(\underline{x}, \underline{y}) \to \exists \underline{z} \forall \underline{w} B_{\mathrm{D}}(\underline{z}, \underline{w})$. In other words, a witness $\underline{x}$ to $\forall \underline{y} A_{\mathrm{D}}(\underline{x}, \underline{y})$ gives rise to a witness $\underline{z}$ to $\forall \underline{w} B_{\mathrm{D}}(\underline{z}, \underline{w})$. If this is done by a rule of computation, there should be finite-type computable functionals $\underline{f}$ such that $\forall \underline{x}(\forall \underline{y} A_{\mathrm{D}}(\underline{x}, \underline{y}) \to \forall \underline{w} B_{\mathrm{D}}(\underline{f}(\underline{x}), \underline{w}))$. Gödel invites us to interprete the inner implication in the following way: Whenever a counter-example $\underline{w}$ is given to $B_{\mathrm{D}}(\underline{f}(\underline{x}),\underline{w})$ then a counter-example $\underline{y}$ is given to $A_{\mathrm{D}}(\underline{x},\underline{y})$. If the latter counter-example is given by a rule of computation in terms of the former, then there should be finite-type computable functionals $\underline{g}$ such that $\neg B_{\mathrm{D}}(\underline{f}(\underline{x}),\underline{w})) \to \neg A_{\mathrm{D}}(\underline{x},\underline{g}(\underline{x},\underline{w}))$. Using the decidability of quantifier-free formulas, we get the implication $A_{\mathrm{D}}(\underline{x},\underline{g}(\underline{x},\underline{w})) \to B_{\mathrm{D}}(\underline{f}(\underline{x}),\underline{w})$, as in the definition. The attentive reader will notice that several of the passages above are not intuitionistically justifiable. However, Gödel's definitions do sustain a soundness theorem. There are three important principles in connection with this theorem:

I. The axiom of choice AC^ω.

II. The principle of independence of premises stated for universal antecedents, denoted by $\mathsf{IP}^\omega_\forall$ (a formula is *universal* if it is of the form $\forall\underline{x}A_{\mathrm{qf}}(\underline{x})$, for A_{qf} quantifier-free).

III. *Markov's Principle* MP^ω: $\neg\forall\underline{z}A_{\mathrm{qf}}(\underline{z}) \to \exists\underline{z}\neg A_{\mathrm{qf}}(\underline{z})$, where A_{qf} is a quantifier-free formula and the z's are of *any* types.

Markov's principle *strictu sensu* applies only for z of type 0. The acceptance of this principle differentiates the school of Russian constructivism from the other constructivist schools. Its acceptance is based on the intuition that if $\forall z^0 A_{\mathrm{qf}}(z)$ leads to a contradiction then, if we test in succession each natural number z^0 for the (decidable) truth of $A_{\mathrm{qf}}(z)$ then we will eventually find a z^0 such that $\neg A_{\mathrm{qf}}(z)$. Note that this intuition does not apply for z of non-zero type. The name "Markov's Principle" for non-zero types is a misnomer.[5]

Using the decidability of quantifier-free formulas, we can see that MP^ω implies $(\forall\underline{z}A_{\mathrm{qf}}(\underline{z}) \to B_{\mathrm{qf}}) \to \exists\underline{z}(A_{\mathrm{qf}}(\underline{z}) \to B_{\mathrm{qf}})$, for A_{qf} and B_{qf} quantifier-free formulas. Note that MP^ω is the particular case when A_{qf} is $\perp$.

Theorem 4.2 (Soundness of the *dialectica* interpretation)**.** *Suppose that*

$$\mathsf{HA}^\omega_0 + \mathsf{AC}^\omega + \mathsf{IP}^\omega_\forall + \mathsf{MP}^\omega + \Delta \vdash A(\underline{z}),$$

where Δ is a set of universal sentences and A is an arbitrary formula (with the free variables as shown). Then there are closed terms $\underline{t}$ of appropriate types such that

$$\mathsf{HA}^\omega_0 + \Delta \vdash \forall\underline{z}\forall\underline{y}A_{\mathrm{D}}(\underline{t}(\underline{z}),\underline{y},\underline{z}).$$

As usual, the proof is by induction on the length of formal derivations, and the terms are effectively constructed from the formal derivations. The discussion of the seemingly innocuous contraction axiom $A \to A \wedge A$ is subtle in three respects (with the advent of *Linear Logic* in the late eighties, we learned that contraction is not that innocuous). Firstly, the choice of witnessing functionals is not canonical

[5]A referee of this paper pointed out that in the model of continuous functionals ECF^ω, where there is a dense subset of finitary objects at any type, it makes some sense to retain the name "Markov's principle."

at this point. Secondly, it involves a definition by cases functional. Finally, it uses the decidabilility of quantifier-free formulas. The interpretation of the hypothesis is $\exists x \forall y A_{\mathrm{D}}(x, y)$ and that of the conclusion is

$$\exists x_1 \exists x_2 \forall y_1 \forall y_2 (A_{\mathrm{D}}(x_1, y_1) \wedge A_{\mathrm{D}}(x_2, y_2)).$$

We must define functionals f_1, f_2 and g such that

$$A_{\mathrm{D}}(x, g(x, y_1, y_2)) \rightarrow A_{\mathrm{D}}(f_1(x), y_1) \wedge A_{\mathrm{D}}(f_2(x), y_2).$$

We take f_1 and f_2 as $\lambda x.x$. Suppose that y_1 and y_2 are of type σ. In order to define g we need a closed term C_σ of type $0 \rightarrow \sigma \rightarrow \sigma \rightarrow \sigma$ which satisfies in HA_0^ω the following "definition by cases" requirements:

$$B[C_\sigma(0, u, v)/w] \leftrightarrow B[u/w] \text{ and } B[C_\sigma(Sz^0, u, v)/w] \leftrightarrow B[v/w],$$

where w is a distinguished variable of type σ and B is a quantifier-free formula. The functional C_σ can be defined with the aid of the recursor R_σ by $R_\sigma(z, u, \lambda w \lambda s.v)$. Note that the checking of the contraction axiom needs a recursor, not merely the combinators. Since A_{D} is quantifier-free, there is a closed term t of $\mathcal{L}_0^\omega$ such that $\mathsf{HA}_0^\omega \vdash t(x, y) = 0 \leftrightarrow \neg A_{\mathrm{D}}(x, y)$. It is easy to see that g can be taken to be $\lambda x \lambda y_1 \lambda y_2.C_\sigma(t(x, y_1), y_1, y_2)$. But, as remarked, the choice is not canonical at this point: we can also take the term $\lambda x \lambda y_1 \lambda y_2.C_\sigma(t(x, y_2), y_2, y_1)$.

The following proposition is an immediate consequence of the soundness theorem:

Proposition 4.3 (Extraction and conservation, *dialectica* case)**.** *Suppose that*

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega + \Delta \vdash \forall x \exists y A_{\mathrm{qf}}(x, y),$$

where Δ is a set of universal sentences and A_{qf} is a quantifier-free formula with free variables among x and y. Then there is a closed term t of appropriate type such that

$$\mathsf{HA}_0^\omega + \Delta \vdash \forall x A_{\mathrm{qf}}(x, tx).$$

Theorem 4.4 (Characterization)**.** *For any formula A,*

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega \vdash A \leftrightarrow (A)^{\mathrm{D}}.$$

With the aid of the characterization theorem, we can prove the following sound extraction result for *arbitrary* formulas A:

Proposition 4.5. *Let A be an* arbitrary *formula with free variables among x and y. Suppose that*

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega \vdash \forall x \exists y A(x, y),$$

Then there is a closed term t of appropriate type such that

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega \vdash \forall x A(x, tx).$$

Proof. Let $(A)^{\mathrm{D}}(x, y)$ be $\exists \underline{z} \forall \underline{w} A_{\mathrm{D}}(y, \underline{z}, \underline{w}, x)$. By the soundness theorem, there are closed terms t and $\underline{q}$ of appropriate types such that

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega \vdash \forall x \forall \underline{w} A_{\mathrm{D}}(tx, \underline{q}x, \underline{w}, x).$$

The result follows by the characterization theorem. □

4.2. **On the monotone functional interpretation.** The reason why we could include universal sentences Δ in the statement of the soundness theorem for the *dialectica* interpretation is because the *dialectica* interpretation of a universal sentence is (essentially) itself. By weakening the conclusion of the soundness theorem it is possible to deal with a wider classe of sentences.

Theorem 4.6 (Soundness of the monotone functional interpretation)**.** *Suppose that*

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega + \Delta \vdash A(\underline{z}),$$

with Δ is a set of sentences *of the form $\exists \underline{x} \leq \underline{r} \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y})$, where B_{qf} is quantifier-free, $\underline{r}$ is tuple of closed terms, and A is an arbitrary formula (with the free variables as shown). Then there are closed monotone terms $\underline{t}$ of appropriate types such that*

$$\mathsf{HA}_0^\omega + \Delta \vdash \exists \underline{x} \leq^* \underline{t}\, \forall \underline{z} \forall \underline{y}\, A_{\mathrm{D}}(\underline{x}(\underline{z}), \underline{y}, \underline{z}).$$

The proof is by induction on the length of formal derivations, and the terms are effectively constructed from the formal derivations. In the proof, when it comes to the contraction axiom, the choice of terms becomes canonical. However, both the decidability of quantifier-free formulas and the definition by cases functional are still required.

Proposition 4.7 (Extraction and conservation, monotone case)**.** *Suppose that*

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega + \Delta \vdash \forall x \exists y A_{\mathrm{qf}}(x, y),$$

with Δ is a set of sentences *of the form $\exists \underline{x} \leq \underline{r} \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y})$, where B_{qf} is quantifier-free, $\underline{r}$ is tuple of closed terms, and A_{qf} is a quantifier-free formula with free variables among x and y. Then there is a closed monotone term t of appropriate type such that*

$$\mathsf{HA}_0^\omega + \Delta \vdash \forall x, u\, (x \leq^* u \to \exists y \leq^* tu\, A_{\mathrm{qf}}(x, y)).$$

When x is of type 1 we can put u as x^{M} and get $\forall x \exists y \leq^* tx A_{\mathrm{qf}}(x, y)$. When furthermore y is of type 0, even $\forall x A_{\mathrm{qf}}(x, tx)$ is in order.

With the aid of the characterization theorem of the *dialectica* interpretation (notice that the assignment of formulas for the *dialectica* and monotone interpretations are the same), the following sound extraction result for *arbitrary* formulas A can be proved in the manner of Proposition 4.5:

Proposition 4.8. *Let A be an* arbitrary *formula with free variables among x and y. Suppose that*

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega \vdash \forall x \exists y A(x, y),$$

Then there is a closed monotone term t of appropriate type such that

$$\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega \vdash \forall x, u\, (x \leq^* u \to \exists y \leq^* tu\, A(x, y)).$$

In order to compare the *dialectica* with the monotone interpretation, let us consider a toy, but illuminating, example: the lesser limited principle of omniscience LLPO (see Section 3.6). This principle is not intuitionistically acceptable. Furthermore, it does not have a *dialectica* interpretation. Let us see why not. Fix two recursively enumerable, recursively inseparable sets X and Y. Consider quantifier-free formulas $A_{\mathrm{qf}}(w,k)$ and $B_{\mathrm{qf}}(w,r)$ such that $X = \{w \in \mathbb{N} : \exists k\, \neg A_{\mathrm{qf}}(w,k)\}$ and $Y = \{w \in \mathbb{N} : \exists r\, \neg B_{\mathrm{qf}}(w,r)\}$. The *dialectica* assignment of the following instance of LLPO (with numerical parameter w)

$$\forall k, r(A_{\mathrm{qf}}(w,k) \vee B_{\mathrm{qf}}(w,k)) \to \forall k A_{\mathrm{qf}}(w,k) \vee \forall r B_{\mathrm{qf}}(w,r)$$

is essentially

$$\exists n^1 \exists f, g \forall w \forall k, r\big(A_{\mathrm{qf}}(w, fkrw) \vee B_{\mathrm{qf}}(w, gkrw) \to (nw = 0 \to A_{\mathrm{qf}}(w,k)) \wedge (nw \neq 0 \to B_{\mathrm{qf}}(w,r))\big).$$

A *dialectica* interpretation would provide a closed term t of type 1 such that

$$\forall w \forall k, r\big((tw = 0 \to A_{\mathrm{qf}}(w,k)) \wedge (tw \neq 0 \to B_{\mathrm{qf}}(w,r))\big),$$

is true (at this juncture, we are using the fact that X and Y are disjoint). But this would entail that the set $\{w \in \mathbb{N} : t^{\mathrm{S}^\omega} w = 0\}$ is recursive and separates X from Y.

Nonetheless, it has a monotone functional interpretation (in a suitable verifying theory). One must find closed terms t and q of types $0 \to 0$ and $0 \to 0 \to 0 \to 0$, respectively, such that

$$(\star)\ \exists n \leq t \exists f, g \leq q \forall w \forall k, r\big(A_{\mathrm{qf}}(w, fwkr) \vee B_{\mathrm{qf}}(w, gwkr) \to (nw = 0 \to A_{\mathrm{qf}}(w,k)) \wedge (nw \neq 0 \to B_{\mathrm{qf}}(w,r))\big).$$

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

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

and verifies that, for each w, $\forall k, r(\overline{A}(w,k) \vee \overline{B}(w,r))$. By LLPO, one gets, for each w, $\forall k \overline{A}(w,k) \vee \forall r \overline{B}(w,r)$. 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. The functions f and g are, respectively, $\lambda w, k, r.(\mu m \leq \max(k,r)\, \neg A_{\mathrm{qf}}(w,m))$ and $\lambda w, k, r.(\mu m \leq \max(k,r)\, \neg B_{\mathrm{qf}}(w,m))$ (where $\mu m \leq t$ is the bounded minimization operation; we take $t+1$ as the default value if no pertinent value satisfies the matrices).

Let us take stock. Suppose that the theory $\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega + \mathsf{LLPO}$ proves $A(\underline{z})$. It is easy to see that $\mathsf{HA}_0^\omega + (\star) \vdash \mathsf{LLPO}$. Therefore, $\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega + (\star) \vdash A(\underline{z})$. Since $(\star)$ has the right syntactic form, by Theorem 4.6 there are closed terms $\underline{t}$ such that $\mathsf{HA}_0^\omega + (\star) \vdash \exists \underline{x} \leq^* \underline{t}\, \forall \underline{z} \forall \underline{y}\, A_{\mathrm{D}}(\underline{x}(\underline{z}), \underline{y}, \underline{z})$. In the paragraphs above we have shown that $(\star)$ is a consequence of HA_0^ω, LLPO, but a bit of choice is also needed (AC^ω is, of course, sufficient). So, in order to verify $\exists \underline{x} \leq^* \underline{t}\, \forall \underline{z} \forall \underline{y}\, A_{\mathrm{D}}(\underline{x}(\underline{z}), \underline{y}, \underline{z})$ we need more than just $\mathsf{HA}_0^\omega + \mathsf{LLPO}$.

The above is a typical phenomenom. Usually, one deals with principles of the form

$$(\star\star)\quad \exists \underline{x} \leq \underline{r}\underline{w} \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y}, \underline{w}),$$

with *parameters* $\underline{w}$ (as axioms, one should of course take the universal closures of these principles). The effect of having parameters is that these principles are no longer covered by Theorem 4.6. One must instead use the *uniformization* of these principles, namely the *sentences* $\exists \underline{X} \leq \underline{r} \forall \underline{w} \forall \underline{y} B_{\mathrm{qf}}(\underline{X}(\underline{w}), \underline{y}, \underline{w})$. Note that these uniformizations have the right syntactical form for the application of Theorem 4.6. In other words, if one wants to state a monotone soundness theorem which includes in Δ principles of the form $(\star\star)$, as Kohlenbach does, then in the verifying theory one must *strengthen* these principles by their uniformizations. As long as one's aim is just to extract true computational information (in the form of bounding terms $\underline{t}$) from proofs, the uniformization procedure is acceptable because if a certain *true* principle of the form $(\star\star)$ is used in the proof then its uniformization is also true. Therefore, the verification of the role of the extracted terms takes place in a true theory and, hence, *correct* computational information is indeed extracted. The pre-eminent example is weak König's lemma (see Section 3.6). In its tree formulation, WKL is the statement

$$\forall T^1(Tree_\infty(T) \to \exists \alpha \leq_1 1\, \forall n^0\, T(\overline{\alpha} n) = 0),$$

where we are using the sequence notation of Section 2.4, and $Tree_\infty(T)$ abbreviates the conjunction of

$$\forall s^0(Ts = 0 \to Seq_2(s)) \wedge \forall s, u(Tu = 0 \wedge s \preceq u \to Ts = 0)$$

with the infinity clause $\forall n^0 \exists s(Ts = 0 \wedge |s| = n)$. Here, $Seq_2(s)$ expresses that s is the number-code of a binary sequence, $s \preceq r$ means that the binary sequence given by s is an initial segment of r, and $|s|$ is the length of the binary sequence given by s. Within HA_0^ω, this principle can be put in the form $(\star\star)$. NB the verification of this takes some work and was done in [26]. Therefore, by the above discussion, one has a monotone soundness theorem with WKL, although with a strengthened version of it in the verifying theory of the soundness theorem. In fact, this strengthened version can be taken to be the so-called *uniform* weak König's lemma:

$$\exists \Phi^{1 \to 1} \forall T^1(Tree_\infty(T) \to \Phi(T) \leq_1 1 \wedge \forall n^0\, T(\overline{\Phi(T)}\, n) = 0).$$

Even though one needs in general to strengthen WKL to the uniform weak König's lemma in the verifying theory, in certain particular cases the *dialectica* interpretation together with some majorizability tricks also permits the *elimination* of WKL from the verifying theory: this happens when $A(\underline{z})$ is an existencial numerical statement with parameters $\underline{z}$ of types 0 or 1.

4.3. **Negative translation.** Let us consider Gödel-Gentzen's negative translation in the framework of arithmetic.

Definition 4.9. *To each formula A of the language $\mathcal{L}_0^\omega$ we associate its (Gödel-Gentzen)* negative translation A^{g} *according to the following clauses:*

i. A^{g} *is A, for atomic formulas A.*
ii. $(A \wedge B)^{\mathrm{g}}$ *is* $A^{\mathrm{g}} \wedge B^{\mathrm{g}}$.
iii. $(A \vee B)^{\mathrm{g}}$ *is* $\neg(\neg A^{\mathrm{g}} \wedge \neg B^{\mathrm{g}})$.
iv. $(A \to B)^{\mathrm{g}}$ *is* $A^{\mathrm{g}} \to B^{\mathrm{g}}$.
v. $(\forall x A)^{\mathrm{g}}$ *is* $\forall x A^{\mathrm{g}}$.
vi. $(\exists x A)^{\mathrm{g}}$ *is* $\neg \forall x \neg A^{\mathrm{g}}$.

In the framework of pure logic we must define A^{g} as $\neg\neg A$ for atomic A, but this is not needed in the current framework because atomic formulas are decidable. The result of Gödel-Gentzen states that if a formula A is classically provable from Γ then A^{g} is intuitionistically provable from Γ^{g}, where $\Gamma^{\mathrm{g}} = \{B^{\mathrm{g}} : B \in \Gamma\}$. This result extends to HA_0^ω because the negative translation of an induction axiom is still an induction axiom and because the remainder axioms are universal formulas. It does *not* extend to $\mathsf{HA}_0^\omega + \mathsf{AC}^\omega + \mathsf{IP}_\forall^\omega + \mathsf{MP}^\omega$ because of the axiom of choice (the other principles pose no problem since they are laws of classical logic). However, the restriction of the axiom of choice for quantifier-free matrices $\mathsf{AC}_{\mathrm{qf}}^\omega$ behaves well under the negative translation *provided that* Markov's principle MP^ω is present in the verifying theory (as well as $\mathsf{AC}_{\mathrm{qf}}^\omega$ itself). In fact:

Theorem 4.10 (Negative translation)**.** *Suppose that*

$$\mathsf{PA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \Delta \vdash A,$$

where Δ is a set of sentences and A is an arbitrary sentence. Then

$$\mathsf{HA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \mathsf{MP}^\omega + \Delta^{\mathrm{g}} \vdash A^{\mathrm{g}}.$$

Let us analyze the negative translation of $\mathsf{AC}_{\mathrm{qf}}^\omega$. It is,

$$\forall x \neg \forall y \neg B_{\mathrm{qf}}(x, y) \to \neg \forall f \neg \forall x B_{\mathrm{qf}}(x, fx),$$

where B_{qf} is a quantifier-free formula. It is easy to check that this translation is provable in $\mathsf{HA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \mathsf{MP}^\omega$. To see this, assume $\forall x \neg \forall y \neg B_{\mathrm{qf}}(x, y)$. By MP^ω, we get $\forall x \exists y B_{\mathrm{qf}}(x, y)$. By $\mathsf{AC}_{\mathrm{qf}}^\omega$, $\exists f \forall x B_{\mathrm{qf}}(x, fx)$ and, *a fortiori*, $\neg \forall f \neg \forall x B_{\mathrm{qf}}(x, fx)$.

4.4. **Extraction and all that (III) & (IV).** We may now state an extraction and conservation result which is applicable to a *classical* theory:

Proposition 4.11 (Extraction and conservation, classical *dialectica* case)**.** *Suppose that*

$$\mathsf{PA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \Delta \vdash \forall x \exists y A_{\mathrm{qf}}(x, y),$$

where Δ is a set of universal sentences and A_{qf} is a quantifier-free formula with free variables among x and y. Then there is a closed term t of appropriate type such that

$$\mathsf{HA}_0^\omega + \Delta \vdash A_{\mathrm{qf}}(x, tx).$$

This is an easy corollary of the properties of the negative translation and the soundness of the *dialectica* interpretation. If $\forall x \exists y A_{\mathrm{qf}}(x, y)$ is provable in $\mathsf{PA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \Delta$ then, by Theorem 4.10, $\forall x \neg\forall y \neg A_{\mathrm{qf}}(x, y)$ is provable in $\mathsf{HA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \mathsf{MP}^\omega + \Delta$. Then so is $\forall x \exists y A_{\mathrm{qf}}(x, y)$ (because MP^ω is available!). At this point, we use Theorem 4.2.

In a similar vein, the combination of the negative translation and the soundness of the monotone functional interpretation yields,

Proposition 4.12 (Extraction and conservation, classical monotone case)**.** *Suppose that*

$$\mathsf{PA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \Delta \vdash \forall x \exists y A_{\mathrm{qf}}(x, y),$$

where Δ is a set of sentences of the form $\exists \underline{x} \leq \underline{r} \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y})$, with B_{qf} quantifier-free, $\underline{r}$ a tuple of closed terms and $\underline{y}$ of any types, and A_{qf} is a quantifier-free formula with free variables among x and y. Then there is a closed monotone term t of appropriate type such that

$$\mathsf{HA}_0^\omega + \Delta \vdash \forall x, u\, (x \leq^* u \to \exists y \leq^* tu\, A_{\mathrm{qf}}(x, y)).$$

Observe that each sentence in Δ proves intuitionistically its own Gödel-Gentzen translation. As in Proposition 4.7, when x is of type 1 we may conclude $\forall x \exists y \leq^* tx A_{\mathrm{qf}}(x, y)$. When furthermore y is of type 0, even $\forall x A_{\mathrm{qf}}(x, tx)$ is in order.

Theorem 4.13 (Characterization)**.** *For any formula A,*

$$\mathsf{PA}_0^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega \vdash A \leftrightarrow (A^{\mathrm{g}})^{\mathrm{D}}.$$

4.5. **The no-counterexample interpretation.** It is a basic observation in pure logic that a first-order formula in prenex normal form

$$\exists x_0 \forall y_0 \exists x_1 \forall y_1 \ldots \exists x_k \forall y_k\, A_{\mathrm{qf}}(x_0, y_0, x_1, y_1, \ldots, x_k, y_k),$$

is classically valid if, and only if, its *Herbrandization*

$$\exists x_0 \exists x_1 \ldots \exists x_k\, A_{\mathrm{qf}}(x_0, f_0(x_0), x_1, f_1(x_0, x_1), \ldots, x_k, f_k(x_0, \ldots, x_k))$$

is classically valid, where $f_0, f_1, \ldots, f_k$ are *new* function symbols of appropriate arities (known as *index functions*). Perhaps the most intuitive way of seeing this is to show that the negation of the former formula is satisfiable if, and only if, the negation of the latter one is.

Theorem 4.14 (No-counterexample interpretation)**.** *Suppose that the sentence*

$$\exists x_0 \forall y_0 \ldots \exists x_k \forall y_k\, A_{\mathrm{qf}}(x_0, y_0, \ldots, x_k, y_k),$$

of the language of first-order arithmetic is provable in PA *(first-order Peano arithmetic). Then there are closed terms* $t_0, \ldots, t_k$ *of such that*

$$\mathsf{HA}_0^\omega \vdash \forall f_0 \ldots \forall f_k\, A_{\mathrm{qf}}(t_0(\underline{f}), f_0(t_0(\underline{f})), \ldots, t_k(\underline{f}), f_k(t_0(\underline{f}), \ldots, t_k(\underline{f}))),$$

where $\underline{f}$ *abbreviates a tuple of variables* $f_0, f_1, \ldots, f_k$ *of (essentially) type 1.*

Observation. A type of the form $0 \to 0 \to \ldots \to 0$ is essentially of type 1 via a pairing code.

Proof. Note that $\exists x_0 \ldots \exists x_k\, A_{\mathrm{qf}}(x_0, f_0(x_0), \ldots, x_k, f_k(x_0, \ldots, x_k))$ follows from $\exists x_0 \forall y_0 \ldots \exists x_k \forall y_k\, A_{\mathrm{qf}}(x_0, y_0, \ldots, x_k, y_k)$ by logic alone. Therefore, the former formula is provable in PA_0^ω. We now use the extraction theorem of the previous section. □

To make sense of the name of the theorem, we may think of $f_0, \ldots, f_k$ as denoting functions in (essentially) S_1 that attempt to provide a counterexample to the truth of $\exists x_0 \forall y_0 \ldots \exists x_k \forall y_k\, A_{\mathrm{qf}}(x_0, y_0, \ldots, x_k, y_k)$, by making $A(n_0, f_0(n_0), \ldots, n_k, f_k(n_0, \ldots, n_k))$ false for any given numerical values $n_0, \ldots, n_k$. However, such counterexample must fail for the values $n_0 = t_0^{\mathrm{S}^\omega}(\underline{f}), \ldots, n_k = t_k^{\mathrm{S}^\omega}(\underline{f})$. As we have argued at the end of Section 2.4, the functions $t_0^{\mathrm{S}^\omega}, \ldots, t_k^{\mathrm{S}^\omega}$ are computable (note that we can view the closed terms $t_0, \ldots, t_k$ as having type 2). Hence, the above theorem says that values which defeat a purported counterexample may be effectively constructed from the attempted counterexample *and* that the effective computations are specified by closed terms of $\mathcal{L}_0^\omega$.

The no-counterexample interpretations coincides with the *dialectica* interpretation (after a double negation translation) for $\forall\exists\forall$ statements. This is the case, for instance, with the modified infinite convergence theorem discussed in the introduction. However, the two interpretations already differ for $\exists\forall\exists$ statements. Kohlenbach discusses in several places the shortcomings of the no-counterexample interpretation vis-à-vis the *dialectica* interpretation. One such discussion can be found in his recent book [35].

4.6. **Digression on provably total functions.** The following result can be proved formalizing Tait's normalization argument:

Proposition 4.15. *Let* $t[x_1, \ldots, x_k]$ *be a term with its (free) variables as shown, all of which are of type 0. The theory* HA_0^ω *proves the* Π_2^0*-sentence saying that for all natural numbers* $n_1, \ldots, n_k$ *the closed term* $t[\overline{n_1}/x_1, \ldots, \overline{n_k}/x_k]$ *normalizes. As a consequence, so does the theory* HA*.*

A word of caution: PA_0^ω does not prove the *sentence* that says that every closed term of $\mathcal{L}_0^\omega$ has a normal form because this would imply that PA_0^ω proves its own consistency, an impossibility by Gödel's second incompleteness theorem. To see this, suppose that PA_0^ω proves '0=1'. By the *proof* of the soundness of the *dialectica* interpretation there would be a sequence of closed terms $\underline{t}_n$ of $\mathcal{L}_0^\omega$ and a sequence of quantifier-free formulas $A_n(\underline{x}, \underline{y})$ ending in '0=1' such that $A_n(\underline{t}_n, \underline{q})$ holds for each list of closed terms $\underline{q}$ of appropriate types. Well, a truth predicate for these (universal) statements can be defined within PA_0^ω *provided that* each closed term normalizes (reducing the verifications of the quantifier-free matrices to checking whether pairs of numerals are, or are not, the same), and *this* would yield a consistency proof.

The following corollary is immediate, but worth an explicit formulation:

Corollary 4.16. *Let us fix a closed term t of type 1. The theory HA_0^ω proves the Π_2^0-sentence saying that, for each numeral $\overline{n}$, the term $t\overline{n}$, has a normal form (necessarily a numeral). As a consequence, so does the theory* HA.

The previous corollary and Proposition 4.11 imply that the provably total Σ_1^0-functions of PA are given by the closed terms of type 1 of $\mathcal{L}_0^\omega$. This provides an alternative characterization to the one based on Gentzen's work on the ordinal analysis of PA ($< \varepsilon_0$-recursion).

Let us now briefly consider a very natural subsystem of HA_0^ω. The theory iPRA^ω differs from HA_0^ω by only having the recursor R_0 and, correspondingly, induction in the following restricted form:

$$A_{\mathrm{qf}}(0) \wedge \forall x^0(A_{\mathrm{qf}}(x) \to A_{\mathrm{qf}}(Sx)) \to \forall x A_{\mathrm{qf}}(x),$$

where A_{qf} is quantifier-free. Due to the absence of higher-order recursors, in order to interprete the contraction axioms we need *primitive* constants C_σ in the language satisfying the "decision by cases" requirements discussed in Section 4.1.

The combination of Gödel's *dialectica* interpretation and the Gödel-Gentzen negative translation yields,

Proposition 4.17 (Extraction and conservation, classical p.r. case)**.** *Suppose that*

$$\mathsf{PRA}^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega + \Delta \vdash \forall x \exists y A_{\mathrm{qf}}(x, y),$$

where Δ is a set of universal sentences and A_{qf} is a quantifier-free formula with free variables among x and y (PRA^ω is the classical *theory associated with iPRA^ω). Then there is a closed term t of appropriate type such that*

$$\mathsf{iPRA}^\omega \vdash A_{\mathrm{qf}}(x, tx).$$

In the present setting, there is a also a notion of contraction of terms and a corresponding strong normalization theorem, with the following consequence:

Proposition 4.18. *For every closed term t of type 1 of the language of iPRA^ω, t^{S^ω} is a primitive recursive function.*

It is easy to see that the theory $\mathsf{PRA}^\omega + \mathsf{AC}_{\mathrm{qf}}^{0,0}$ and, *a fortiori* $\mathsf{PRA}^\omega + \mathsf{AC}_{\mathrm{qf}}^\omega$, proves the following two schemes:

(1) Σ_1^0-*induction*: $A(0) \wedge \forall x^0(A(x) \to A(Sx)) \to \forall x A(x)$, for Σ_1^0-formulas A, i.e., formulas of the form $\exists n^0 A_{\mathrm{qf}}(n, x)$ with A_{qf} quantifier-free.

(2) Δ^0_1-*comprehension*: $\forall x^0(A(x) \leftrightarrow \neg B(x)) \rightarrow \exists \alpha^1 \forall x^0(\alpha x = 0 \leftrightarrow A(x))$, where A and B are Σ^0_1-formulas.

The previous schemes are the distinguished axioms of the second-order theory RCA_0, which plays an important role in the studies of *Reverse Mathematics*. The above results show that the witnesses of the Π^0_2-consequences of $\mathsf{RCA_0}$ are the primitive recursive functions. *A fortiori*, they are also the witnesses of the Π^0_2-consequences of the subsystem PA^1 of PA, characterized by having the scheme of induction restricted to Σ^0_1-formulas (this result is, essentially, due to Charles Parsons and, independently, to Grigori Mints and Gaisi Takeuti).

4.7. **Suggested reading and historical notes.** Gödel's functional interpretation was published in German in [17]. An English translation of this paper can be found in Gödel's collected works [19]. The soundness theorem, generalized to finite-type Heyting arithmetic, with the characteristic principles, is implicit in [40]. The characterization theorem is due to Mariko Yasugi [56]. The books [55] and [35] present proofs of these theorems. [2] is a good survey of the functional interpretations until the mid-nineties. The monotone functional interpretation is due to Kohlenbach and appeared in [27]. A good reference for this interpretation is [35]. The treatment of LLPO in Section 4.2 is related to, but not quite the same as in [31]. The elimination of WKL mentioned at the end of Section 4.2 appears in [24]. Similar eliminations are also explained in [2] and [35]. An alternative elimination of WKL via the elimination of LLPO is worked out in [31].

The negative translation is due, independently, to Gödel and Gerhard Gentzen (cf. [16]). The characterization theorem of Section 4.4 is essentially due to Kreisel in [40]. Instead of dealing with classical Peano arithmetic via the negative translation followed by the *dialectica* interpretation, a direct and elegant path is provided by Joseph Shoenfield in [45] (note, however, that Shoenfield's interpretation is the same as the combination of an appropriate negative translation followed by the *dialectica* interpretation: this was recently shown by Kohlenbach and Thomas Streicher in [38]). The no-counterexample interpretation is due to Kreisel in [39]; [30] discusses in detail the shortcomings of the no-counterexample interpretation. Proofs of the results in Section 4.6 can be found in [55]. The result of Charles Parsons appeared in [44]. For information regarding the program of Reverse Mathematics one should consult [46].

5. Injecting uniformities

5.1. **Intensional majorizability.** Bounded formulas are treated as computationally empty by the bounded modified realizability interpretation, in the sense that their realizers are trivial. This hinges on the fact that the majorizability relations of

Howard-Bezem are given by $\tilde{\exists}$-free formulas and that these formulas have empty realizers. However, the functional interpretation acts non-trivially on these formulas. We solve this problem with the notion of *intensional* majorizability.

We introduce a modification of the language $\mathcal{L}^\omega_\leq$, dubbed $\mathcal{L}^\omega_\trianglelefteq$. The new language is an extension of $\mathcal{L}^\omega_0$ with the primitive binary relation symbol $\leq$ (as in $\mathcal{L}^\omega_\leq$), but also with *primitive* binary relation symbols $\trianglelefteq_\sigma$, for each type σ. Note that the terms of $\mathcal{L}^\omega_\trianglelefteq$ and $\mathcal{L}^\omega_0$ are the same. The symbols $\trianglelefteq_\sigma$ are the intensional counterparts of $\leq^*_\sigma$. There are now new atomic formulas, and the syntactic device of bounded quantification is modified in such a way that it now concerns the intensional symbols instead of the extensional ones. I.e., we have quantifications of the form $\forall x \trianglelefteq t A(x)$ and $\exists x \trianglelefteq t A(x)$, for terms t not containing x. In the current framework, we use the terminology of the bounded formulas and monotone functionals *with respect to the intensional symbols.*

Definition 5.1. *The theory* $\mathsf{HA}^\omega_\trianglelefteq$ *is an altered version of* $\mathsf{HA}^\omega_\leq$ *with the axiom schemes:*

$$\mathsf{B}_\forall \ : \ \forall x \trianglelefteq t A(x) \leftrightarrow \forall x (x \trianglelefteq t \to A(x)),$$
$$\mathsf{B}_\exists \ : \ \exists x \trianglelefteq t A(x) \leftrightarrow \exists x (x \trianglelefteq t \wedge A(x))$$

instead of the corresponding extensional ones, and with the further axioms

$$\mathsf{M}_1 \ : \ x \trianglelefteq_0 y \leftrightarrow x \leq y,$$
$$\mathsf{M}_2 \ : \ x \trianglelefteq_{\rho\to\sigma} y \to \forall u \trianglelefteq_\rho v (xu \trianglelefteq_\sigma yv \wedge yu \trianglelefteq_\sigma yv)$$

and a rule $\mathsf{RL}_\trianglelefteq$

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

where s and t are terms of $\mathcal{L}^\omega_\trianglelefteq$, A_{bd} is a bounded formula and u and v are variables that do not occur free in the conclusion. Moreover, the induction scheme is now extended to all formulas of the language $\mathcal{L}^\omega_\trianglelefteq$.

We observe that it is not needed to state the axioms for combinators, the axiom of equality and the axioms for recursors for the new atomic formulas of the language, since these already follow from the more restricted statements.

The crucial feature of the above theory is that we have *rules* instead of the purported axioms: $\forall u \trianglelefteq_\rho v(xu \trianglelefteq_\sigma yv \wedge yu \trianglelefteq_\sigma yv) \to x \trianglelefteq_{\rho\to\sigma} y$. Note that the previous formula would pose a problem for the functional interpretation. The rules, however, do *not* pose such a problem! We dubbed the new majorizability symbols $\trianglelefteq_\sigma$ *intensional* because they are (partially) governed by *rules*. The presence of rules entails the failure of the deduction theorem in $\mathsf{HA}^\omega_\trianglelefteq$, a feature that many do not find attractive. However, one should keep in mind that the rules are introduced for *mathematical reasons* and that "mathematical attraction" is partly a question of familiarity. In the end, systems where rules play an essential role must be judged by their own mathematical merits.

Even though the rules are *weaker* than the axioms we still have:

Lemma 5.2. *The theory* $\mathsf{HA}^{\omega}_{\trianglelefteq}$ *proves that the relations* $\trianglelefteq_\sigma$ *are transitive and that* $x \trianglelefteq y \to y \trianglelefteq y$. *For type 1, we have* $x \trianglelefteq_1 y \to x \leq_1^* y$, $x \trianglelefteq_1 x^{\mathrm{M}}$ *and* $\min_1(x,y) \trianglelefteq_1 y^{\mathrm{M}}$ *(where the* $\min_1$ *function is the minimum function defined pointwise). Howard's majorizability theorem holds: For each closed term t, there is a closed term q such that* $\mathsf{HA}^{\omega}_{\trianglelefteq} \vdash t \trianglelefteq q$. *Furthermore, it holds of the very same term constructed for the extensional case.*

5.2. **Bounded functional interpretation.** Let us now define the bounded functional interpretation:

Definition 5.3. *To each formula A of the language* $\mathcal{L}^{\omega}_{\trianglelefteq}$ *we assign formulas* $(A)^{\mathrm{B}}$ *and* A_{B} *so that* $(A)^{\mathrm{B}}$ *is of the form* $\tilde{\exists}\underline{b}\tilde{\forall}\underline{c}A_{\mathrm{B}}(\underline{b},\underline{c})$, *with* $A_{\mathrm{B}}(\underline{b},\underline{c})$ *a bounded formula, according to the following clauses:*

1. $(A)^{\mathrm{B}}$ *and* A_{B} *are simply A, for atomic formulas A.*

If we have already interpretations for A and B given by $\tilde{\exists}\underline{b}\tilde{\forall}\underline{c}A_{\mathrm{B}}(\underline{b},\underline{c})$ *and* $\tilde{\exists}\underline{d}\tilde{\forall}\underline{e}B_{\mathrm{B}}(\underline{d},\underline{e})$ *(respectively) then we define:*

2. $(A \wedge B)^{\mathrm{B}}$ *is* $\tilde{\exists}\underline{b},\underline{d}\tilde{\forall}\underline{c},\underline{e}(A_{\mathrm{B}}(\underline{b},\underline{c}) \wedge B_{\mathrm{B}}(\underline{d},\underline{e}))$,
3. $(A \vee B)^{\mathrm{B}}$ *is* $\tilde{\exists}\underline{b},\underline{d}\tilde{\forall}\underline{c},\underline{e}(\tilde{\forall}\underline{c'} \trianglelefteq \underline{c}A_{\mathrm{B}}(\underline{b},\underline{c'}) \vee \tilde{\forall}\underline{e'} \trianglelefteq \underline{e}B_{\mathrm{B}}(\underline{d},\underline{e'}))$,
4. $(A \to B)^{\mathrm{B}}$ *is* $\tilde{\exists}\underline{f},\underline{g}\tilde{\forall}\underline{b},\underline{e}(\tilde{\forall}\underline{c} \trianglelefteq \underline{gbe}A_{\mathrm{B}}(\underline{b},\underline{c}) \to B_{\mathrm{B}}(\underline{fb},\underline{e}))$.

For bounded quantifiers we have:

5. $(\forall x \trianglelefteq tA(x))^{\mathrm{B}}$ *is* $\tilde{\exists}\underline{b}\tilde{\forall}\underline{c}\forall x \trianglelefteq t\, A_{\mathrm{B}}(\underline{b},\underline{c},x)$,
6. $(\exists x \trianglelefteq tA(x))^{\mathrm{B}}$ *is* $\tilde{\exists}\underline{b}\tilde{\forall}\underline{c}\exists x \trianglelefteq t\tilde{\forall}\underline{c'} \trianglelefteq \underline{c}\, A_{\mathrm{B}}(\underline{b},\underline{c'},x)$.

And for unbounded quantifiers we define:

7. $(\forall x A(x))^{\mathrm{B}}$ *is* $\tilde{\exists}\underline{f}\tilde{\forall}a,\underline{c}\forall x \trianglelefteq aA_{\mathrm{B}}(\underline{f}a,\underline{c},x)$.
8. $(\exists x A(x))^{\mathrm{B}}$ *is* $\tilde{\exists}a,\underline{b}\tilde{\forall}\underline{c}\exists x \trianglelefteq a\tilde{\forall}\underline{c'} \trianglelefteq \underline{c}A_{\mathrm{B}}(\underline{b},\underline{c'},x)$.

There are five important principles in connection with this interpretation:

I. *Intensional Bounded Choice* $\mathsf{bAC}^{\omega}_{\trianglelefteq}$:

$$\forall \underline{x}\exists \underline{y}A(\underline{x},\underline{y}) \to \tilde{\exists}\underline{f}\tilde{\forall}\underline{b}\forall\underline{x} \trianglelefteq \underline{b}\exists\underline{y} \trianglelefteq \underline{f}\underline{b}\, A(\underline{x},\underline{y}),$$

where A is an arbitrary formula of $\mathcal{L}^{\omega}_{\trianglelefteq}$. It is clear that the forms of choice $\mathsf{bAC}^{i,j}_{\mathrm{qf}}$, $\forall x^i \exists y^j B_{\mathrm{qf}}(x,y) \to \exists f \forall x \exists y \leq_j fx\, B_{\mathrm{qf}}(x,y)$, for $i,j \in \{0,1\}$ and B_{qf} quantifier-free, follow from $\mathsf{bAC}^{\omega}_{\trianglelefteq}$ in $\mathsf{HA}^{\omega}_{\trianglelefteq}$ (even if $\mathsf{bAC}^{\omega}_{\trianglelefteq}$ were restricted to bounded matrices only, a principle which we denote by $\mathsf{bBAC}^{\omega}_{\trianglelefteq}$). Due to the availability of minimization for quantifier-free formulas, note that $\mathsf{bAC}^{i,0}_{\mathrm{qf}} \Rightarrow \mathsf{AC}^{i,0}_{\mathrm{qf}}$, for $i \in \{0,1\}$. Here, $\mathsf{AC}^{i,j}_{\mathrm{qf}}$, $i,j \in \{0,1\}$, is the usual form of choice $\forall x^i \exists y^j B_{\mathrm{qf}}(x,y) \to \exists f \forall x B_{\mathrm{qf}}(x,fx)$, with

B_{qf} quantifier-free. Note that appropriate tuple versions of these principles also follow from $\mathsf{bBAC}^{\omega}_{\trianglelefteq}$.

II. *Intensional Bounded Independence of Premises* $\mathsf{bIP}^{\omega}_{\trianglelefteq}$:

$$(A \to \exists \underline{y} B(\underline{y})) \to \tilde{\exists} \underline{b}(A \to \exists \underline{y} \trianglelefteq \underline{b}\, B(\underline{y})),$$

where A is a universal formula (with bounded matrix) and B is an arbitrary formula. In the present setting, by "universal with bounded matrix" we mean a formula of the form $\forall \underline{x} A_{\mathrm{bd}}(\underline{x})$, with A_{bd} a *bounded (intensional) matrix.*

III. *Intensional Bounded Markov's Principle* $\mathsf{MP}^{\omega}_{\trianglelefteq}$:

$$(\forall \underline{y} A_{\mathrm{bd}}(\underline{y}) \to B_{\mathrm{bd}}) \to \tilde{\exists} \underline{b}(\forall \underline{y} \trianglelefteq \underline{b} A_{\mathrm{bd}}(\underline{y}) \to B_{\mathrm{bd}}),$$

where A_{bd} is a bounded matrix and B_{bd} is a bounded formula. When B is the formula $0 = 1$, the above principle specializes to $\neg \forall \underline{y} A_{\mathrm{bd}}(\underline{y}) \to \tilde{\exists} \underline{b} \neg \forall \underline{y} \trianglelefteq \underline{b} A_{\mathrm{bd}}(\underline{y})$. If y is of type 0 and the matrix is quantifier-free, the principle further specializes to the familiar Markov's principle of type 0: $\neg \forall y A_{\mathrm{qf}}(y) \to \exists y \neg A_{\mathrm{qf}}(y)$. In this, we are using bounded numerical search.

IV. *Intensional Bounded Contra-Collection Principle* $\mathsf{bBCC}^{\omega}_{\trianglelefteq}$:

$$\tilde{\forall} \underline{b} \exists \underline{z} \trianglelefteq \underline{c} \forall \underline{y} \trianglelefteq \underline{b} A_{\mathrm{bd}}(\underline{y}, \underline{z}) \to \exists \underline{z} \trianglelefteq \underline{c} \forall \underline{y} A_{\mathrm{bd}}(\underline{y}, \underline{z}),$$

where $\underline{c}$ is a tuple of monotone functionals and A_{bd} is a bounded formula. Note that this is *classically* equivalent to collection restricted to bounded matrices (see Proposition 5.4 below). This principle allows the conclusion of certain existentially bounded statements from the assumption of weakenings thereof. Let us see that it implies weak König's lemma WKL (the statement of WKL is in the end of Section 4.2). Suppose that $Tree_{\infty}(T^1)$. The infinity clause says that $\forall n^0 \exists s (Ts = 0 \wedge |s| = n)$. We introduce some notation. Given s a binary sequence, let $\hat{s}$ be the infinite binary path which extends s by appending an infinite string of zeros. With the aid of the rule $\mathsf{RL}_{\trianglelefteq}$ one can show that $\mathsf{HA}^{\omega}_{0} \vdash \forall s (Seq_2(s) \to \hat{s} \trianglelefteq_1 1)$. Therefore, from the infinity clause one gets $\forall n^0 \exists \alpha \trianglelefteq_1 1\, (T(\overline{\alpha} n) = 0)$. Of course, by the definition of tree,

$$\forall n^0 \exists \alpha \trianglelefteq_1 1 \forall k \leq n\, (T(\overline{\alpha} k) = 0).$$

Applying $\mathsf{bBCC}^{\omega}_{\trianglelefteq}$, we may infer $\exists \alpha \trianglelefteq_1 \forall n (T(\overline{\alpha} n) = 0)$. Since $\alpha \trianglelefteq_1 1 \to \alpha \leq 1$, we get our infinite path through T.

The following principle, dubbed *Intensional Bounded Disjunction Property*, also follows from $\mathsf{bBCC}^{\omega}_{\trianglelefteq}$:

$$\tilde{\forall} \underline{b} \tilde{\forall} \underline{c} (\forall \underline{x} \trianglelefteq \underline{b} A_{\mathrm{bd}}(\underline{x}) \vee \forall \underline{y} \trianglelefteq \underline{c} B_{\mathrm{bd}}(\underline{y})) \to \forall \underline{x} A_{\mathrm{bd}}(\underline{x}) \vee \forall \underline{y} B_{\mathrm{bd}}(\underline{y}),$$

where A_{bd} and B_{bd} are bounded formulas. This property clearly implies (within $\mathsf{HA}^{\omega}_{\trianglelefteq}$) the lesser limited principle of omniscience LLPO, mentioned in Section 3.6.

V. *Intensional Majorizability Axioms* $\mathsf{MAJ}^{\omega}_{\trianglelefteq}$: $\forall x \exists y (x \trianglelefteq y)$.

The following result is similar to Proposition 3.13:

Proposition 5.4. *The theory* $\mathsf{HA}^\omega_{\trianglelefteq}+\mathsf{bAC}^\omega_{\trianglelefteq}+\mathsf{bIP}^\omega_{\trianglelefteq}$ *proves the* Intensional Collection Principle $\mathsf{bC}^\omega_{\trianglelefteq}$*:*

$$\forall \underline{z} \trianglelefteq \underline{c}\exists \underline{y} A(\underline{y},\underline{z}) \to \tilde{\exists} \underline{b} \forall \underline{z} \trianglelefteq \underline{c} \exists \underline{y} \trianglelefteq \underline{b}\, A(\underline{x},\underline{y}),$$

where A *is an arbitrary formula and* $\underline{c}$ *is a tuple of monotone functionals.*

Obviously, the restricted theory $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bBAC}^\omega_{\trianglelefteq} + \mathsf{bIP}^\omega_{\trianglelefteq}$ proves the Intensional *Bounded* Collection Scheme $\mathsf{bBC}^\omega_{\trianglelefteq}$, that is, the above scheme restricted to bounded formulas A.

The theory $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega_{\trianglelefteq} + \mathsf{bIP}^\omega_{\trianglelefteq}$ is *classically inconsistent.* E.g., it refutes the *limited principle of omniscience* LPO (in Errett Bishop's terminology):

$$\forall x^1 (\forall n^0 (xn = 1) \vee \exists n^0 (xn \neq 1)).$$

To see this, assume the above. Hence $\forall x^1 \exists n^0 (\forall k^0 (xk = 1) \vee xn \neq 1)$, by intuitionistic logic. *A fortiori*, $\forall x \trianglelefteq_1 1 \exists n^0 (\forall k (xk = 1) \vee xn \neq 1)$. By the intensional collection principle $\mathsf{bC}^\omega_{\trianglelefteq}$, we get $\exists m^0 \forall x \trianglelefteq_1 1 \exists n \leq m (\forall k (xk = 1) \vee xn \neq 1)$. Take such m and consider the sequence s of length $m+1$ with constant value 1. Clearly, $\hat{s} \trianglelefteq_1 1$ but it is not the case that $\exists n \leq m (\forall k (\hat{s}k = 1) \vee \hat{s}n \neq 1)$.

Also, the restricted theory $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bBAC}^\omega_{\trianglelefteq} + \mathsf{bIP}^\omega_{\trianglelefteq} + \mathsf{MP}^\omega_{\trianglelefteq}$ already refutes a basic form of extensionality. In fact, it proves the *negation* of

$$\forall \Phi^2 \forall \alpha^1, \beta^1 \, (\forall k^0 (\alpha k = \beta k) \to \Phi\alpha = \Phi\beta).$$

Towards a contradiction, assume the above. In particular, one has

$$\forall \Phi \trianglelefteq_2 1^2 \, \forall \alpha, y \trianglelefteq_1 1^1 \exists k \, (\alpha k = \beta k \to \Phi\alpha = \Phi\beta),$$

where $1^1 := \lambda k^0.1^0$ and $1^2 := \lambda \gamma^1.1^1$ (we used here $\mathsf{MP}^\omega_{\trianglelefteq}$). By $\mathsf{bBC}^\omega_{\trianglelefteq}$, one may infer

$$\exists n \forall \Phi \trianglelefteq_2 1 \forall \alpha, y \trianglelefteq_1 1 \, (\forall k < n (\alpha k = \beta k) \to \Phi\alpha = \Phi\beta).$$

Take one such $n = n_0$. Define Φ according to:

$$\gamma^1 \leadsto_\Phi \begin{cases} 0 & \text{if } \forall k \leq n_0 \, (\gamma k = 0) \\ 1 & \text{otherwise} \end{cases}$$

It is clear that for $\alpha := \lambda k.0$ and $\beta := \lambda k.\delta_{n_0,k}$ (Kronecker's delta) one has $\forall k < n_0 \, (\alpha k = \beta k)$ but $\Phi\alpha \neq \Phi\beta$. Since it is easy to show that $\Phi \trianglelefteq 1^2$ and $\alpha, \beta \trianglelefteq 1^1$, we are faced with a contradiction.

The above two examples are not set-theoretically sound. The bounded functional interpretation *injects uniformities* which are absent in the universe of sets and which are incompatible with it. Given this state of affairs, it is pressing to assure that the theory $\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega_{\trianglelefteq} + \mathsf{bIP}^\omega_{\trianglelefteq} + \mathsf{MP}^\omega_{\trianglelefteq} + \mathsf{bBCC}^\omega_{\trianglelefteq} + \mathsf{MAJ}^\omega_{\trianglelefteq}$ is consistent.

Theorem 5.5 (Soundness of the bounded functional interpretation)**.** *Suppose that*

$$\mathsf{HA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega_{\trianglelefteq} + \mathsf{bIP}^\omega_{\trianglelefteq} + \mathsf{MP}^\omega_{\trianglelefteq} + \mathsf{bBCC}^\omega_{\trianglelefteq} + \mathsf{MAJ}^\omega_{\trianglelefteq} + \Delta \vdash A(\underline{z}),$$

where Δ is a set of universal sentences (with bounded matrices) and A is an arbitrary formula (with the free variables as shown). Then there are closed monotone terms of appropriate types such that

$$\mathsf{HA}^\omega_\trianglelefteq + \Delta \vdash \tilde{\forall} \underline{a} \forall \underline{z} \trianglelefteq \underline{a} \tilde{\forall} \underline{c}\, A_{\mathrm{B}}(\underline{t}(\underline{a}), \underline{c}, \underline{z}).$$

Given an appropriate restatement, we may include in Δ sentences of the form $\exists \underline{x} \leq \underline{r} \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y})$, as in the soundness theorem of the monotone functional interpretation (see Theorem 4.6). Let us see what we mean by this. Given a tuple of closed terms $\underline{r}$, by Lemma 5.2 there is a tuple of closed terms $\underline{t}$ such that $\mathsf{HA}^\omega_\trianglelefteq \vdash \underline{r} \trianglelefteq \underline{t}$. The sentence $\exists \underline{x} \trianglelefteq \underline{t}(\underline{x} \leq \underline{r} \wedge \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y}))$ obviously implies the original sentence $\exists \underline{x} \leq \underline{r} \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y})$. Furthermore, the modified sentence is of the form $\exists \underline{x} \trianglelefteq \underline{t}\, \forall \underline{w} C_{\mathrm{qf}}(\underline{x}, \underline{w})$ for quantifier-free C_{qf}, because the formula $\underline{x} \leq \underline{r}$ is universal. By $\mathsf{bBCC}^\omega_\trianglelefteq$, the modified sentence is implied by $\tilde{\forall} \underline{b} \exists \underline{x} \trianglelefteq \underline{t}\, \forall \underline{w} \trianglelefteq \underline{b}\, C_{\mathrm{qf}}(\underline{x}, \underline{w})$, and this latter sentence has the right form for applying the soundness theorem above. Of course, in the conclusion of the soundness theorem, the verifying theory must include this sentence. At this juncture, we draw attention to the fact that this sentence "flattens" (see Section 5.5) to a sentence which is implied by $\forall \underline{u} \exists \underline{x} \leq \underline{r} \forall \underline{y} \leq^* \underline{u}\, B_{\mathrm{qf}}(\underline{x}, \underline{y})$ (we are using here (iii) of Lemma 2.4). NB this is a *weaker* statement than the original one. This should be compared with the monotone functional interpretation (see the discussion in Section 4.2) where a strengthening of the original statement is needed.

As usual, the proof of the above soundness theorem is by induction on the length of formal derivations, and the terms are effectively constructed from the formal derivations. However, for the bounded functional interpretation the choice of the witnessing terms is *always canonical*, and there is no need for a definition by cases functional nor for the decidability of bounded formulas.

Proposition 5.6 (Extraction and conservation, intuitionistic intensional case). *Suppose that*

$$\mathsf{HA}^\omega_\trianglelefteq + \mathsf{bAC}^\omega_\trianglelefteq + \mathsf{bIP}^\omega_\trianglelefteq + \mathsf{MP}^\omega_\trianglelefteq + \mathsf{bBCC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq + \Delta \vdash \forall x \exists y A_{\mathrm{bd}}(x, y),$$

where Δ is a set of universal sentences (with bounded matrices) and A_{bd} is a bounded formula with free variables among x and y. Then there is a closed monotone term t of appropriate type such that

$$\mathsf{HA}^\omega_\trianglelefteq + \Delta \vdash \tilde{\forall} a \forall x \trianglelefteq a \exists y \trianglelefteq ta\, A_{\mathrm{bd}}(x, y).$$

Theorem 5.7 (Characterization). *For any formula A,*

$$\mathsf{HA}^\omega_\trianglelefteq + \mathsf{bAC}^\omega_\trianglelefteq + \mathsf{bIP}^\omega_\trianglelefteq + \mathsf{MP}^\omega_\trianglelefteq + \mathsf{bBCC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq \vdash A \leftrightarrow (A)^{\mathrm{B}}.$$

Using the above characterization theorem and an argument as in the proof of Proposition 4.5, we get the following fact:

Proposition 5.8. *Let A be an* arbitrary *formula with free variables among x and y. Suppose that*

$$\mathsf{HA}^{\omega}_{\trianglelefteq} + \mathsf{bAC}^{\omega}_{\trianglelefteq} + \mathsf{bIP}^{\omega}_{\trianglelefteq} + \mathsf{MP}^{\omega}_{\trianglelefteq} + \mathsf{bBCC}^{\omega}_{\trianglelefteq} + \mathsf{MAJ}^{\omega}_{\trianglelefteq} \vdash \forall x \exists y A(x,y),$$

Then there is a closed monotone term t of appropriate type such that

$$\mathsf{HA}^{\omega}_{\trianglelefteq} + \mathsf{bAC}^{\omega}_{\trianglelefteq} + \mathsf{bIP}^{\omega}_{\trianglelefteq} + \mathsf{MP}^{\omega}_{\trianglelefteq} + \mathsf{bBCC}^{\omega}_{\trianglelefteq} + \mathsf{MAJ}^{\omega}_{\trianglelefteq} \vdash \tilde{\forall} a \forall x \trianglelefteq a \exists y \trianglelefteq ta\, A(x,y).$$

It is not clear what is accomplished by the above extraction result, given that the verifying theory is not sound. This is in sharp contrast with Proposition 4.8, where the extraction is indeed sound. We are here facing the same phenomenon as the one discussed at the end of Section 3.5. However, note that for restricted A, Proposition 5.6 yields a sound extraction result since the characteristic principles are absent from the verifying theory (the forthcoming Section 5.5 is also relevant for this discussion).

In the presence of classical logic, the situation concerning the monotone and the bounded interpretations changes. In the classical case, the matrix A must necessarily be narrowed down (it is well-known that computational extraction in the classical case may already fail for universal matrices). The reader should compare the extraction results of Proposition 4.12 and Proposition 5.15 below.

5.3. **Digression on a new conservation result.** In Section 4.6, we mentioned the classical second-order theory $\mathsf{RCA_0}$, the ordinary base theory for Reverse Mathematics. By $\mathsf{iRCA_0}$ we mean an intuitionistic version of this theory, whereby the logic used is intuitionistic, adjoined with the axiom $\forall X \forall x (x \in X \vee x \notin X)$. Let us introduce some further second-order principles. $\mathsf{bIP}_{\forall}$ is the following *bounded* version of the scheme of independence of premises:

$$(A \to \exists y B(y)) \to \exists y (A \to \exists w \leq y B(w)),$$

where A is a formula starting with a string of universal (first or second-order) quantifiers followed by a formula without unbounded first-order quantifications (it can have bounded first-order quantifications as well as second-order quantifications), and B is an arbitrary formula. MP is the usual (numerical) Markov's principle. $\mathsf{bAC}^{\mathbb{N}}$ is *bounded* version of the countable axiom of choice:

$$\forall x \exists y A(x,y) \to \exists X \big(\mathrm{Func}(X) \wedge \forall x, y (\langle x, y\rangle \in X \to \exists w \leq y A(x,w))\big),$$

where A is any formula, and $\mathrm{Func}(X)$ says that the set X is constituted by the codes of the pairs of a total function from $\mathbb{N}$ to $\mathbb{N}$. Finally, FAN is the scheme

$$\forall X \exists x A(x,X) \to \exists z \forall X \exists x \leq z A(x,X),$$

with arbitrary A. Weak König's lemma WKL has its usual formulation in terms of infinite binary trees.

Theorem 5.9. *The second-order intuitionistic theory*

$$\mathsf{iRCA_0} + \mathsf{WKL} + \mathsf{MP} + \mathsf{bIP}_{\forall} + \mathsf{bAC}^{\mathbb{N}} + \mathsf{FAN}$$

is conservative over PA^1 *with respect to* Π^0_2*-sentences.*[6]

Proof. The second-order language of arithmetic can be embedded in $\mathcal{L}^\omega_{\trianglelefteq}$ by letting the first-order variables run over type 0 arguments, letting the second-order variables run over type 1 variables X such that $X \trianglelefteq_1 1$, and by interpreting $x \in X$ by $Xx = 0$. Under this embedding, the theory $\mathsf{iRCA_0} + \mathsf{WKL} + \mathsf{MP} + \mathsf{bIP}_\forall + \mathsf{bAC}^{\mathbb{N}} + \mathsf{FAN}$ is clearly a sub-theory of $\mathsf{iPRA}^\omega_{\trianglelefteq} + \mathsf{bAC}^\omega_{\trianglelefteq} + \mathsf{bIP}^\omega_{\trianglelefteq} + \mathsf{MP}^\omega_{\trianglelefteq} + \mathsf{bBCC}^\omega_{\trianglelefteq} + \mathsf{MAJ}^\omega_{\trianglelefteq}$ (where $\mathsf{iPRA}^\omega_{\trianglelefteq}$ is the obvious intensional version of iPRA^ω). Therefore, if the former theory proves a Π^0_2-sentence then, by (an adaptation of) Proposition 5.6, then $\mathsf{PRA}^\omega_{\trianglelefteq}$ already proves it. By *flattening* (see the forthcoming Section 5.5), so does the theory PRA^ω. By an internalization argument in the manner of Proposition 2.3, it can be argued that PRA^ω is (first-order) conservative over PA^1. We are done. □

As a consequence, the provably total Σ^0_1-functions of $\mathsf{iRCA_0} + \mathsf{WKL} + \mathsf{MP} + \mathsf{bIP}_\forall + \mathsf{bAC}^{\mathbb{N}} + \mathsf{FAN}$ are the primitive recursive functions (even though this theory is *classically* inconsistent; see the argument after Proposition 5.4). In Reverse Mathematics, $\mathsf{WKL_0}$ is the classical theory $\mathsf{RCA_0}$ together with weak König's lemma WKL. The importance of the theory $\mathsf{WKL_0}$ is well documented from the work in Reverse Mathematics. The following elimination result is originally due to Harvey Friedman.

Corollary 5.10. *The theory* $\mathsf{WKL_0}$ *is* Π^0_2*-conservative over* $\mathsf{RCA_0}$.

Proof. Suppose that $\mathsf{WKL_0}$ proves a certain Π^0_2-sentence. Note that $(\mathsf{WKL_0})^g$ is a subtheory of $\mathsf{iRCA_0} + \mathsf{WKL} + \mathsf{MP}$. Therefore, by the negative translation, $\mathsf{iRCA_0} + \mathsf{WKL} + \mathsf{MP}$ also proves the given Π^0_2-sentence (notice the presence of MP). The result follows from the previous theorem. □

5.4. **Extraction and all that (V).** The negative translation of Gödel-Gentzen can be easily extended to the language $\mathcal{L}^\omega_{\trianglelefteq}$ according to the extra clauses:

vii. $(\forall x \trianglelefteq t\, A)^{\mathrm{g}}$ *is* $\forall x \trianglelefteq t\, A^{\mathrm{g}}$.
viii. $(\exists x \trianglelefteq t\, A)^{\mathrm{g}}$ *is* $\neg \forall x \trianglelefteq t\, \neg A^{\mathrm{g}}$.

It is clear that the negative translation of a bounded formula is still bounded, and it can be shown by induction on the type that the intensional majorizability relations are *stable*, i.e., $\mathsf{HA}^\omega_{\trianglelefteq} \vdash \neg\neg(x \trianglelefteq y) \to x \trianglelefteq y$. Now, in analogy with the *dialectica* interpretation, the next result is not difficult:

Theorem 5.11 (Negative translation)**.** *Suppose that*

[6] There is a careless misstep in our abstract [9], where the scheme of independence of premisses $\mathsf{IP}_\forall$ (i.e., $(A \to \exists y B(y)) \to \exists y (A \to B(y))$, for A and B as in $\mathsf{bIP}_\forall$) and the principle of countable choice $\mathsf{AC}^{\mathbb{N}}$ (i.e., $\forall x \exists y A(x,y) \to \exists X[\mathrm{Func}(X) \wedge \forall x, y(\langle x, y\rangle \in X \to A(x,y))]$) appear instead of the weaker principles $\mathsf{bIP}_\forall$ and $\mathsf{bAC}^{\mathbb{N}}$, respectively. We do not know if the theorem still holds with $\mathsf{IP}_\forall$ and $\mathsf{AC}^{\mathbb{N}}$.

$$\mathsf{PA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq + \Delta \vdash A,$$

where Δ *is a set of sentences,* A *is an arbitrary sentence and* $\mathsf{PA}^\omega_\trianglelefteq$ *is the classical version of* $\mathsf{HA}^\omega_\trianglelefteq$*. Then,*

$$\mathsf{HA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq + \mathsf{MP}^\omega_\trianglelefteq + \Delta^{\mathrm{g}} \vdash A^{\mathrm{g}}.$$

Remember that $\mathsf{PA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq$ proves the intensional collection principle restricted to bounded formulas $\mathsf{bBC}^\omega_\trianglelefteq$ and, by *classical logic*, the contra-collection scheme $\mathsf{bBCC}^\omega_\trianglelefteq$ follows. The next proposition is a consequence of the combination of the negative translation with the soundness theorem:

Theorem 5.12 (Extraction and conservation, classical intensional case)**.** *Suppose that*

$$\mathsf{PA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq + \Delta \vdash \forall x \exists y A_{\mathrm{bd}}(x,y),$$

where Δ *is a set of universal sentences (with bounded matrices) and* A_{bd} *is a bounded formula with free variables among* x *and* y*. Then there is a closed monotone term* t *of appropriate type such that*

$$\mathsf{PA}^\omega_\trianglelefteq + \Delta \vdash \tilde{\forall} a \forall x \trianglelefteq a \exists y \trianglelefteq ta\, A_{\mathrm{bd}}(x,y).$$

In contrast with the analogous result of Section 4.3, the verification theory is a *classical* theory because intensional bounded formulas are not decidable in general. However, if y is of type 0 and A_{bd} is quantifier-free, the verification can be done in $\mathsf{HA}^\omega_\trianglelefteq + \Delta^{\mathrm{g}}$.

Theorem 5.13 (Characterization)**.** *For any formula* A*,*

$$\mathsf{PA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq \vdash A \leftrightarrow (A^{\mathrm{g}})^{\mathrm{B}}.$$

5.5. **Flattening.** We want to use this *intensional* technology in "real world" applications. The following lemma is the passageway from the intensional theory to plain $\mathsf{PA}^\omega_\leq$:

Lemma 5.14 (Flattening)**.** *Suppose* $\mathsf{PA}^\omega_\trianglelefteq + \Gamma \vdash A$*, where* A *is a sentence and* Γ *is a set of sentences, formulated in the intensional language* $\mathcal{L}^\omega_\trianglelefteq$*. Then* $\mathsf{PA}^\omega_\leq + \Gamma^* \vdash A^*$*, where* B^* *is the sentence of* $\mathcal{L}^\omega_\leq$ *obtained from* B *by replacing throughout the binary symbols* $\trianglelefteq_\sigma$ *by the formulas* $\leq^*_\sigma$ *(*mutatis mutandis *for sets of sentences).*

We call B^* the *flattening* of B (*mutatis mutandis* for sets of sentences).

Proposition 5.15 (Extraction and conservation, classical flattened case)**.** *Suppose that*

$$\mathsf{PA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq + \Delta \vdash \forall x \exists y A_{\mathrm{bd}}(x,y),$$

where Δ *is a set of universal sentences (with bounded intensional matrices) and* A_{bd} *is a bounded (intensional) formula with free variables among* x *and* y*. Then there is a closed monotone term* t *of appropriate type such that*

$$\mathsf{PA}^{\omega}_{\leq} + \Delta^* \vdash \forall a \forall x \leq^* a \exists y \leq^* ta\, A^*_{\mathrm{bd}}(x,y).$$

In particular, the theory $\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{bBAC}^{\omega}_{\trianglelefteq} + \mathsf{MAJ}^{\omega}_{\trianglelefteq}$ is conservative over $\mathsf{PA}^{\omega}_{\leq}$ with respect to Π^0_2-sentences.

It is interesting to inquire what happens if one flattens the theory $\mathsf{PA}^{\omega}_{\trianglelefteq}$ together with the characteristic principles of Theorem 5.13. It so happens that $\mathsf{PA}^{\omega}_{\leq} + \mathsf{bBAC}^{\omega}$ is *inconsistent* (here bBAC^{ω} is the flattening of $\mathsf{bBAC}^{\omega}_{\trianglelefteq}$). Let us see why. Well, $\mathsf{PA}^{\omega}_{\trianglelefteq}$ proves

$$\forall x^1 (x \leq^*_1 0 \to x \trianglelefteq_1 0) \to \forall x \trianglelefteq_1 1\, \exists n^0\, (\neg x \trianglelefteq_1 0 \to xn \neq 0).$$

If we now apply the intensional collection principle restricted to bounded formulas to the consequent above, we get

$$\forall x^1 (x \leq^*_1 0 \to x \trianglelefteq_1 0) \to \exists m^0 \forall x \trianglelefteq_1 1 \exists n \leq m(\neg x \trianglelefteq_1 0 \to xn \neq 0).$$

We remind the reader that the intensional collection principle restricted to bounded formulas is provable in $\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{bBAC}^{\omega}_{\trianglelefteq}$. Hence, by flattening, $\mathsf{PA}^{\omega}_{\leq} + \mathsf{bBAC}^{\omega}$ proves

$$\forall x^1 (x \leq^*_1 0 \to x \leq^*_1 0) \to \exists m^0 \forall x \leq^*_1 1 \exists n \leq m(\neg x \leq^*_1 0 \to xn \neq 0).$$

Since the antecedent is logically true, we infer

$$\exists m^0 \forall x \leq 1 (\exists n (xn \neq 0) \to \exists n \leq m\, (xn \neq 0)),$$

a statement which is obviously refuted in $\mathsf{PA}^{\omega}_{\leq}$. Therefore, the theory $\mathsf{PA}^{\omega}_{\leq} + \mathsf{bBAC}^{\omega}$ is inconsistent.

5.6. **On extensionality and uniform boundedness.** We say that a formula $A(x^1)$ is *extensional* in the type 1 variable x if $\forall x^1, y^1\, (x =_1 y \wedge A(x) \to A(y))$, and we write $Ext_x[A]$.

Proposition 5.16. *Let $A_{\mathrm{bd}}(x^1, k^0)$ be a bounded (intensional) formula. Then the theory $\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{bBAC}^{\omega}_{\trianglelefteq} + \mathsf{MAJ}^{\omega}_{\trianglelefteq}$ proves the implication whose antecedent is $Ext_x[A_{\mathrm{bd}}]$ and whose consequent is*

$$\forall x \leq_1 z \exists k A_{\mathrm{bd}}(x,k) \to \exists n \forall x \leq_1 z \exists k \leq n A_{\mathrm{bd}}(x,k).$$

Proof. Assume $\forall x \leq_1 z \exists k A_{\mathrm{bd}}(x,k)$. By the extensionality of A_{bd} with respect to x, we have $\forall x \exists k A_{\mathrm{bd}}(\min_1(x,z),k)$. *A fortiori*, we get $\forall x \trianglelefteq z^{\mathrm{M}} \exists k A_{\mathrm{bd}}(\min_1(x,z),k)$. Hence, by the intensional bounded collection scheme $\mathsf{bBC}^{\omega}_{\trianglelefteq}$, we may infer that there is n such that $\forall x \trianglelefteq z^{\mathrm{M}} \exists k \leq n A_{\mathrm{bd}}(\min_1(x,z),k)$. Since $\min_1(x,z) \trianglelefteq z^{\mathrm{M}}$, we have $\forall x \exists k \leq n A_{\mathrm{bd}}(\min_1(\min_1(x,z),z),k)$. By extensionality again, we conclude $\forall x \leq_1 z \exists k \leq n A_{\mathrm{bd}}(x,k)$. □

The principle of the previous proposition is a uniform boundedness principle. As they were stated originally by Kohlenbach, they also incorporate a bit of choice. The following corollary can be proved similarly to the above result with the aid of $\mathsf{bBAC}^{\omega}_{\trianglelefteq}$:

Corollary 5.17. *Let $A_{\mathrm{bd}}(x^1, z^{0\to 1}, m^0, k^0)$ be a bounded (intensional) formula. Then the theory $\mathsf{PA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq$ proves the implication whose antecedent is $Ext_x[A_{\mathrm{bd}}]$ and whose consequent is*

$$\forall m \forall x \leq_1 zm \exists k A_{\mathrm{bd}}(x, z, m, k) \to \exists f^1 \forall m \forall x \leq zm \exists k \leq fm A_{\mathrm{bd}}(x, z, m, k).$$

When the formula A_{bd} is a Σ^0_1-formula (possibly with higher order parameters), the consequent above is dubbed Σ^0_1-UB. Note that the previous corollary includes this case because one can collapse two existential numerical quantifiers into a single one. Observe also that Σ^0_1-UB is a *false* principle. For instance, it entails that *all* type 2 functionals are bounded on compact sets:

$$\forall \Phi^2 \forall z^1 \exists k^0 \forall x \leq_1 z\, (\Phi x \leq k).$$

Full extensionality (see the definition at the end of Section 2.2) cannot be added to the theory $\mathsf{PA}^\omega_\trianglelefteq + \mathsf{bBAC}^\omega_\trianglelefteq + \mathsf{MAJ}^\omega_\trianglelefteq$: as we saw in Section 5.2, this would entail a contradiction. Instead, we consider the theory $\mathsf{PA}^\omega_\leq + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \mathsf{AC}^{0,1}_{\mathrm{qf}} + \Sigma^0_1$-$\mathsf{UB}$ (stated in the language $\mathcal{L}^\omega_\leq$). We will see that adding full extensionality to this theory is consistent (and more). In order to accomplish this, we briefly review a (streamlined version of the) method of elimination of extensionality due to Horst Luckhardt:

Definition 5.18. *We define, by recursion on the type:*

a) $x \approx_0 y \equiv x = y$
b) $x \approx_{\rho\to\tau} y \equiv \forall u^\rho, v^\rho (u \approx_\rho v \to xu \approx_\tau yv)$

Lemma 5.19. *For each finite type σ, the theory HA^ω_0 proves:*

(i) $x \approx_\sigma y \to y \approx_\sigma x$;
(ii) $x \approx_\sigma y \to y \approx_\sigma y$;
(iii) $x \approx_\sigma y \wedge y \approx_\sigma z \to x \approx_\sigma z$;
(iv) $x =_\sigma y \wedge y \approx_\sigma z \to x \approx z$.

Proof. All the claims can be proved by induction on the complexity of the type σ, but (ii) and (iii) should be proved simultaneously. □

Let the expression $E(x)$ abbreviate $x \approx x$ and, given a formula A of $\mathcal{L}^\omega_\leq$, let A^E be the *relativization* of A to the predicate E. Note that $\forall x^1 E(x)$.

Theorem 5.20 (Elimination of extensionality)**.** *Suppose that*

$$\mathsf{E}\text{-}\mathsf{PA}^\omega_\leq + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \Delta \vdash A(\underline{z}),$$

where Δ is a set of universal closures of formulas with bound variables of type 0 or 1 only, and A is an arbitrary formula with its free variables as shown (all this is stated in the language $\mathcal{L}^\omega_\leq$). Then,

$$\mathsf{PA}^\omega_\leq + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \Delta \vdash E(\underline{z}) \to A^E(\underline{z}).$$

Usually, this theorem is also stated with the inclusion of $\mathsf{AC}^{0,1}_{\mathrm{qf}}$. There is however no restriction because this form of choice can be included in Δ (the type $0 \to (0 \to 0)$ is essentially of type 1). Even though $\mathsf{AC}^{1,0}_{\mathrm{qf}}$ is not of the form Δ, the elimination of extensionality still goes through because the type 2 witness functional of $\mathsf{AC}^{1,0}_{\mathrm{qf}}$ can be taken as giving the *least* numerical witness satisfying the matrix of choice, and this forces the functional to satisfy the predicate E (if the parameters of the matrix also do). The following result is due to Kohlenbach:

Theorem 5.21 (Extraction and conservation, uniform boundedness)**.** *Suppose that*

$$\mathsf{E\text{-}PA}^{\omega}_{\leq} + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \mathsf{AC}^{0,1}_{\mathrm{qf}} + \Sigma^0_1\text{-}\mathsf{UB} + \Delta \vdash \forall x^{0/1} \exists y A_{\mathrm{qf}}(x, y),$$

where A_{qf} is a quantifier-free formula with free variables among x and y, and Δ is a set of universal *sentences (all this is stated in the language $\mathcal{L}^{\omega}_{\leq}$). Then there is a closed monotone term t of appropriate type such that*

$$\mathsf{PA}^{\omega}_{\leq} + \Delta \vdash \forall x^{0/1} \exists y \leq^{*} tx\, A_{\mathrm{qf}}(x, y).$$

Proof. For this proof, let us introduce the scheme $(\Sigma^0_1\text{-}\mathsf{UB})^-$ constituted by (universal closures of) implications of the following form: the antecedent is $Ext_x[F]$ and the consequent is $\forall m \forall x \leq_1 zm \exists k F(x, z, m, k) \to \exists f^1 \forall m \forall x \leq zm \exists k \leq fmF(x, z, m, k)$, where $F(x^1, z^{0\to 1}, m^0, k^0)$ is a Σ^0_1-formula (possibly with higher order parameters). Observe that this is a scheme of formulas whose bound variables are (essentially) of type 0 and 1 only. By hypothesis (notice the presence of full extensionality),

$$\mathsf{E\text{-}PA}^{\omega}_{\leq} + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \mathsf{AC}^{0,1}_{\mathrm{qf}} + (\Sigma^0_1\text{-}\mathsf{UB})^- + \Delta \vdash \forall x^{0/1} \exists y A_{\mathrm{qf}}(x, y).$$

Hence, by elimination of extensionality (previous theorem),

$$\mathsf{PA}^{\omega}_{\leq} + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \mathsf{AC}^{0,1}_{\mathrm{qf}} + (\Sigma^0_1\text{-}\mathsf{UB})^- + \Delta \vdash \forall x^{0/1} \exists y A_{\mathrm{qf}}(x, y).$$

By the comments in Section 5.2 apropos intensional bounded choice and by Corollary 5.17, we get $\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{bBAC}^{\omega}_{\trianglelefteq} + \mathsf{MAJ}^{\omega}_{\trianglelefteq} + \Delta \vdash \forall x^{0/1} \exists y A_{\mathrm{qf}}(x, y)$. Hence, according to Theorem 5.12, there is a closed term t of appropriate type such that

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \Delta \vdash \forall x^{0/1} \exists y \trianglelefteq tx\, A_{\mathrm{qf}}(x, y).$$

The desired result follows by flattening. □

The above result can be refined in two ways. On the one hand, Δ may be constituted by sentences of the form $\exists \underline{x} \leq \underline{r} \forall \underline{y} B_{\mathrm{qf}}(\underline{x}, \underline{y})$, with B_{qf} quantifier-free, $\underline{r}$ a tuple of closed terms of type 0 or 1 (the same types of $\underline{x}$) and $\underline{y}$ a tuple of any types. One can apply Luckhardt's elimination of extensionality technique to such sentences and, afterwards, use the techniques discussed after Theorem 5.5. As observed in that discussion, the verification can even be done with the weaker (corresponding) sentences $\forall \underline{z} \exists \underline{x} \leq \underline{r} \forall \underline{y} \leq^{*} \underline{z} B_{\mathrm{qf}}(\underline{x}, \underline{y})$. On the other hand, the above result can also be stated with the form of choice $\mathsf{AC}^{1,1}_{\mathrm{qf}}$. We discuss in detail the latter improvement. This improvement follows from

Proposition 5.22. *The theory* $\mathsf{E\text{-}PA}^\omega_0 + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \Sigma^0_1\text{-}\mathsf{UB}$ *proves* $\mathsf{AC}^{1,1}_{\mathrm{qf}}$.

Proof. We need a preliminary lemma:

Lemma 5.23. *Let* $A_{\mathrm{qf}}(y^1)$ *be a quantifier-free formula. Then*

$$\mathsf{E\text{-}PA}^\omega_0 + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \Sigma^0_1\text{-}\mathsf{UB} \vdash \forall y^1 (A_{\mathrm{qf}}(y) \to \exists m^0 A_{\mathrm{qf}}(\widehat{\overline{y}m})).$$

Proof of the lemma. By extensionality, we have $\forall y, z\,(y =_1 z \wedge A_{\mathrm{qf}}(y) \to A_{\mathrm{qf}}(z))$. Therefore, $\forall y, z \exists n\,(yn = zn \wedge A_{\mathrm{qf}}(y) \to A_{\mathrm{qf}}(z))$. It is easy to see that $\mathsf{AC}^{1,0}_{\mathrm{qf}}$ entails the existence of a functional Σ of type $1 \to (1 \to 0)$ such that

$$\forall y, z (y(\Sigma yz) = z(\Sigma yz) \wedge A_{\mathrm{qf}}(y) \to A_{\mathrm{qf}}(z)).$$

Fix y. By Σ^0_1-UB, there is m^0 such that $\forall z \leq y (\Sigma yz < m)$. The lemma follows. □(of lemma)

We now prove that $\mathsf{AC}^{1,1}_{\mathrm{qf}}$ follows from $\mathsf{E\text{-}PA}^\omega_0 + \mathsf{AC}^{1,0}_{\mathrm{qf}} + \Sigma^0_1\text{-}\mathsf{UB}$. Let $A_{\mathrm{qf}}(x^1, y^1)$ be a quantifier-free formula and suppose that $\forall x \exists y A_{\mathrm{qf}}(x, y)$. By the previous lemma, we may infer that $\forall x \exists s^0 (Seq(s) \wedge A_{\mathrm{qf}}(x, \hat{s}))$, where $Seq(s)$ means that s is a finite sequence of natural numbers. Therefore, by $\mathsf{AC}^{1,0}_{\mathrm{qf}}$, there is a functional Φ of type 2 such that $\forall x (Seq(\Phi x) \wedge A_{\mathrm{qf}}(x, \widehat{\Phi x}))$. Clearly, $\Psi^{1\to 1} := \lambda x.\widehat{\Phi x}$ is a choice function. □(of proposition)

As a consequence of the results above, we obtain an extraction result for the fully extensional, *true* theory, $\mathsf{E\text{-}PA}^\omega_0 + \mathsf{AC}^{1,1}_{\mathrm{qf}}$. However, this result is obtained in a very roundabout way, via a false extension. Is there a more direct route?

5.7. **Suggested reading and historical notes.** The bounded functional interpretation appeared in [12], where the proofs of the theorems in sections 5.1 and 5.2 can be found. The intuitionistic application in Section 5.3 appears here for the first time. The negative translation within the setting of the bounded functional interpretation is also discussed in [12]. A direct interpretation of Peano arithmetic – in the style of Shoenfield – was recently defined in [10]. In this paper, the characterization theorem is formulated with different, but equivalent, characteristic principles. Notwithstanding, Theorem 5.13 regards an indirect interpretation, via a negative translation. However, one could (for instance) use the factorization of Jaime Gaspar [14] to get the result in the text (albeit for the so-called Krivine's negative translation; it is easy to see, though, that the result also holds for the Gödel-Gentzen translation using the fact that these negative translations are intuitionistically equivalent). Flattening is already introduced in [12], but only in [13] it is given its name. The latter article includes a study of the elimination of weak König's lemma in the feasible setting (the elimination technique mentioned at the end of Section 4.2 does not apply to the feasible setting since it uses bounded search in an essential manner). The discussion on extensionality and the uniform boundedness principles in Section 5.6 are based on [12], where stronger results are proved. Luckhardt's result on the elimination of extensionality is from [42] but, as noticed, it is here simplified. The extraction and conservation result on uniform boundedness is due to Kohlenbach and appears (essentially) in [35] under a treatment which is a combination of results in [28] and [32]. Our treatment is rather different, in that it is based on the bounded functional interpretation.

6. Coda

The emphasis of this paper is on the theoretical aspects of proof mining in the style of Kohlenbach and his co-workers. This means studying proof interpretations where the concept of majorizability plays a central role. Even though the applied work has been revolving around *functional* interpretations, for the sake of rounding up we also included in our discussions the realizability interpretations. The main theoretical tool of the applied work has been the monotone functional interpretation but, from a theoretical point of view, we believe that the bounded functional interpretations provide a fresh perspective. We also took the chance of making some comparisons between the monotone and bounded interpretations. It remains to be seen whether the latter interpretations prove to be useful in the applied work. Proof mining itself was left out. I suggest the surveys [37] and [34] for a first reading (see also [36] for a systematic list of statements of the results obtained until 2006). Kohlenbach's book [35] is recommended for a detailed treatment. In [8] the reader can find some *general* discussions on the theoretical and practical benefits of (functional) majorizability interpretations.

We left out two important theoretical topics, one old, the other quite recent. The old one is Clifford Spector's *deep* generalization of Gödel's interpretation to second-order classical arithmetic using bar-recursive functionals (see [47]). The systems which Kohlenbach and his co-workers use in proof mining include, as a matter of course, full second-order comprehension. Kohlenbach's book [35] is a good source for an exposition of Spector's interpretation. Very recently, it was shown in [7] that the bar-recursive functionals of Spector can also be used to obtain a bounded functional interpretation of second-order classical artithmetic. In the words of Part 5, it is possible to inject uniformities into systems containing full second-order comprehension (see [8] for a brief discussion on how far one can go on in doing this). The other topic is the generalization of the monotone functional interpretation to new base types, typically metric or normed spaces. This generalization was introduced in [33] and is also treated in [35]. It has been proved very useful in the applied work and it is rather illuminating from a theoretical point of view.

References

[1] Y. Akama, S. Berardi, S. Hayashi, and U. Kohlenbach. An arithmetical hierarchy of the law of excluded middle and related principles. In *Proceedings of the 19th Annual IEEE Symposium on Logic and Computer Science*, volume 117, pages 192–201. IEEE Press, 2004.

[2] 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.

[3] J. Avigad and H. Towsner. *Functional interpretation and inductive definition*, 2008. arXive:0802.1938v2 [math.LO].

[4] 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.

[5] W. Burr. A Diller-Nahm-style functional interpretation of $\mathsf{KP}\omega$. *Archive for Mathematical Logic*, 39:599–604, 2000.

[6] J. Diller and W. Nahm. Eine Variante zur Dialectica-Interpretation der Heyting-Arithmetik endlicher Typen. *Archive für mathematische Logik und Grundlagenforschung*, 16:49–66, 1974.

[7] P. Engrácia and F. Ferreira. The bounded functional interpretation of the double negation shift. Manuscript.

[8] F. Ferreira. A most artistic package of a jumble of ideas. *dialectica*, 62:205–222, 2008. Special Issue: Gödel's *dialectica* Interpretation. Guest editor: Thomas Strahm.

[9] F. Ferreira. On a new functional interpretation (abstract). *The Bulletin of Symbolic Logic*, 14:128–129, 2008.

[10] F. Ferreira. Injecting uniformities into Peano arithmetic. *Annals of Pure and Applied Logic*, 157:122–129, 2009. Special Issue: Kurt Gödel Centenary Research Prize Fellowships. Editors: Sergei Artemov, Matthias Baaz and Harvey Friedman.

[11] F. Ferreira and A. Nunes. Bounded modified realizability. *The Journal of Symbolic Logic*, 71:329–346, 2006.

[12] F. Ferreira and P. Oliva. Bounded functional interpretation. *Annals of Pure and Applied Logic*, 135:73–112, 2005.

[13] F. Ferreira and P. Oliva. Bounded functional interpretation and feasible analysis. *Annals of Pure and Applied Logic*, 145:115–129, 2005.

[14] J. Gaspar. Factorization of the Shoenfield-like bounded functional interpretation. *Notre Dame Journal of Formal Logic*, 50:53–60, 2009.

[15] P. Gerhardy and U. Kohlenbach. General logical metatheorems for functional analysis. *Transactions of the American Mathematical Society*, 360:2615–2660, 2008.

[16] K. Gödel. Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines Mathematischen Kolloquiums*, 4:34–38, 1933.

[17] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *dialectica*, 12:280–287, 1958. Reprinted with an English translation in [19].

[18] K. Gödel. On an extension of finitary mathematics which has not yet been used, 1972. Published in [19], pages 271–280. Revised version of [17].

[19] K. Gödel. *Collected Works, Vol. II.* S. Feferman *et al.*, eds. Oxford University Press, Oxford, 1990.

[20] W. A. Howard. Hereditarily majorizable functionals of finite type. In A. S. Troelstra, editor, *Metamathematical investigation of intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*, pages 454–461. Springer, Berlin, 1973.

[21] S. C. Kleene. On the interpretation of intuitionistic number theory. *The Journal of Symbolic Logic*, 10:109–124, 1945.

[22] S. C. Kleene. Countable functionals. In A. Heyting, editor, *Constructivity in Mathematics*, pages 81–100. North Holland, Amsterdam, 1959.
[23] U. Kohlenbach. *Theorie der majorisierbaren und stetigen Funktionale und ihre Anwendung bei der Extraktion von Schranken aus inkonstruktiven Beweisen: Effektive Eindeutigkeitsmodule bei besten Approximationen aus ineffektiven Beweisen*. PhD thesis, Frankfurt, pp. xxii+278, 1990.
[24] U. Kohlenbach. Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. *The Journal of Symbolic Logic*, 57:1239–1273, 1992.
[25] U. Kohlenbach. Pointwise hereditary majorization and some applications. *Archive for Mathematical Logic*, 31:227–241, 1992.
[26] 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.
[27] U. Kohlenbach. Analysing proofs in analysis. In W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, editors, *Logic: from Foundations to Applications*, pages 225–260. European Logic Colloquium (Keele, 1993), Oxford University Press, 1996.
[28] U. Kohlenbach. Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals. *Archive for Mathematical Logic*, 36:31–71, 1996.
[29] U. Kohlenbach. Relative constructivity. *The Journal of Symbolic Logic*, 63:1218–1238, 1998.
[30] U. Kohlenbach. On the no-counterexample interpretation. *The Journal of Symbolic Logic*, 64:1491–1511, 1999.
[31] U. Kohlenbach. Intuitionistic choice and restricted classical logic. *Mathematical Logic Quarterly*, 47:455–460, 2001.
[32] U. Kohlenbach. Foundational and mathematical uses of higher types. In W. Sieg et al., editors, *Reflections on the Foundations of Mathematics: Essay in Honor of Solomon Feferman*, volume 15 of *Lecture Notes in Logic*, pages 92–116. A. K. Peters, Ltd., 2002.
[33] U. Kohlenbach. Some logical metatheorems with applications in functional analysis. *Transactions of the American Mathematical Society*, 357:89–128, 2005.
[34] U. Kohlenbach. Proof interpretations and the computational content of proofs in mathematics. *Bulletin of the EATCS*, 93:143–173, 2007.
[35] U. Kohlenbach. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer Monographs in Mathematics. Springer, Berlin, 2008.
[36] U. Kohlenbach. Effective bounds from proofs in abstract functional analysis. In S. B. Cooper, B. Löwe, and A. Sorbi, editors, *New Computational Paradigms: Changing Conceptions of What is Computable*. Springer Verlag, to appear.
[37] 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.
[38] U. Kohlenbach and T. Streicher. Shoenfield is Gödel after Krivine. *Mathematical Logic Quarterly*, 53:176–179, 2007.
[39] G. Kreisel. On the interpretation of non-finitist proofs, part I. *The Journal of Symbolic Logic*, 16:241–267, 1951.
[40] 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.
[41] G. Kreisel. On weak completeness of intuitionistic predicate logic. *The Journal of Symbolic Logic*, 27:139–158, 1962.
[42] H. Luckhardt. *Extensional Gödel Functional Interpretation: A Consistency Proof of Classical Analysis*, volume 306 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973.

[43] P. Oliva. Unifying functional interpretations. *Notre Dame Journal of Formal Logic*, 47:263–290, 2006.
[44] C. Parsons. On a number theoretic choice schema and its relation to induction. In A. Kino J. Myhill and R. E. Vesley, editors, *Intuitionism and Proof Theory*, pages 459–473. North-Holland, 1970.
[45] J. R. Shoenfield. *Mathematical Logic*. Addison-Wesley Publishing Company, 1967. Republished in 2001 by AK Peters.
[46] S. G. Simpson. *Subsystems of Second Order Arithmetic*. Perspectives in Mathematical Logic. Springer, Berlin, 1999.
[47] 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: Proceedings of Symposia in Pure Mathematics*, volume 5, pages 1–27. American Mathematical Society, Providence, Rhode Island, 1962.
[48] T. Strahm (ed.). Gödel's *dialectica* Interpretation. *dialectica*, 62:145–290, 2008.
[49] W. Tait. Intentional interpretations of functionals of finite type I. *The Journal of Symbolic Logic*, 32:198–212, 1967.
[50] T. Tao. The correspondence principle and finitary ergodic theory. Essay posted August 30, 2008. Available at: http://terrytao.wordpress.com/2008/08/30/the-correspondence-principle-and-finitary-ergodic-theory/.
[51] T. Tao. Soft analysis, hard analysis, and the finite convergence principle. In *Structure and Randomness: Pages from Year One of a Mathematical Blog*, pages 17–29. American Mathematical Society, 2008.
[52] A. S. Troelstra. Introductory note to [17] and [18], 1990. Published in [19], pages 214–241.
[53] A. S. Troelstra. Realizability. In S. R. Buss, editor, *Handbook of Proof Theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 408–473. North Holland, Amsterdam, 1998.
[54] A. S. Troelstra and D. van Dalen. *Constructivism in Mathematics. An Introduction*, volume 121 of *Studies in Logic and the Foundations of Mathematics*. North Holland, Amsterdam, 1988.
[55] A. S. Troelstra (ed.). *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973.
[56] M. Yasugi. Intuitionistic analysis and Gödel's interpretation. *Journal of the Mathematical Society of Japan*, 15:101–112, 1963.

Departamento de Matemática, Universidade de Lisboa, Campo Grande, Edifício C6, P-1749-016 Lisboa, Portugal
E-mail address: `ferferr@cii.fc.ul.pt`