TO BE OR NOT TO BE CONSTRUCTIVE

THAT IS NOT THE QUESTION

SAM SANDERS

ABSTRACT. In the early twentieth century, L.E.J. Brouwer pioneered a new philosophy of mathematics, called *intuitionism*. Intuitionism was revolutionary in many respects but stands out -mathematically speaking- for its challenge of *Hilbert's formalist philosophy* of mathematics and rejection of the *law of excluded middle* from the 'classical' logic used in mainstream mathematics. Out of intuitionism grew *intuitionistic logic* and the associated *Brouwer-Heyting-Kolmogorov* interpretation by which 'there exists x' intuitively means 'an algorithm to compute x is given'. A number of schools of constructive mathematics were developed, inspired by Brouwer's intuitionism and invariably based on intuitionistic logic, but with varying interpretations of what constitutes an algorithm. This paper deals with the *dichotomy* between constructive and non-constructive mathematics, or rather the *absence* of such an 'excluded middle'. In particular, we challenge the 'binary' view that mathematics is either constructive or not. To this end, we identify a part of *classical* mathematics, namely *classical Nonstandard Analysis*, and show it inhabits the twilight-zone between the constructive and non-constructive. Intuitively, the predicate 'x is standard' typical of Nonstandard Analysis can be interpreted as 'x is computable', giving rise to computable (and sometimes constructive) mathematics obtained directly from *classical* Nonstandard Analysis. Our results formalise Osswald's longstanding conjecture that classical Nonstandard Analysis is *locally constructive*. Finally, an alternative explanation of our results is provided by Brouwer's thesis that *logic depends upon mathematics*.

1. Introduction

This volume is dedicated to the founder of *intuitionism*, L.E.J. Brouwer, who pursued this revolutionary program with great passion and against his time's received view of mathematics and its foundations ([18, 19, 41, 99, 102]). We therefore find it fitting that our paper attempts to subvert (part of) *our time's* received view of mathematics and its foundations. As suggested by the title, we wish to challenge the binary distinction *constructive*[1] *versus non-constructive* mathematics. We shall assume basic familiarity with constructive mathematics and *intuitionistic logic* with its *Brouwer-Heyting-Kolmogorov* interpretation. A detailed introduction to Nonstandard Analysis is provided in Section 2.

Surprising as this may be to the outsider, the quest for the (ultimate) foundations of mathematics was and is an *ongoing and often highly emotional affair*. The

MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY, LMU MUNICH, GERMANY & DEPARTMENT OF MATHEMATICS, GHENT UNIVERSITY

E-mail address: `sasander@me.com`.

[1]The noun 'constructive' is often used as a synonym for 'effective', while it refers to the foundational framework *constructive mathematics* in logic and the foundations of mathematics ([98]). Context determines the meaning of 'constructive' in this paper (usually the latter).

Grundlagenstreit between Hilbert and Brouwer is perhaps the textbook example (See e.g. [102, II.13]) of a fierce struggle between competing views on the foundations of mathematics, namely Hilbert's formalism and Brouwer's intuitionism. Einstein was apparently disturbed by this controversy and exclaimed the following:

> What is this frog and mouse battle among the mathematicians? ([32, p. 133])

More recently, Bishop mercilessly attacked *Nonstandard Analysis* in his review [13] of Keisler's monograph [53], even going as far as debasing Nonstandard Analysis[2] to a *debasement of meaning* in [14]. Bishop, like Brouwer, believed that to state the existence of an object, one has to provide a construction for it, while Nonstandard Analysis cheerfully includes ideal/non-constructive objects *at the fundamental level*, the textbook example being infinitesimals. Note that Brouwer's student, the intuitionist Arend Heyting, had a higher opinion of Nonstandard Analysis ([42]).

A lot of ink has been spilled over the aforementioned struggles, and we do not wish to add to that literature. By contrast, the previous paragraph is merely meant to establish the well-known juxtaposition of *classical/mainstream/non-constructive* versus *constructive* mathematics[3]. The following quote by Bishop emphasises this 'two poles' view for the specific case of Nonstandard Analysis, which Bishop believed to be the worst exponent of classical mathematics.

> [Constructive mathematics and Nonstandard Analysis] are at opposite poles. Constructivism is an attempt to deepen the meaning of mathematics; non-standard analysis, an attempt to dilute it further. ([12, p. 1-2])

To be absolutely clear, lest we be misunderstood, we only wish to point out the current state-of-affairs in contemporary mathematics: On one hand, there is mainstream mathematics with its classical logic and other fundamentally non-constructive features, of which Nonstandard Analysis is the *nec plus ultra* according to some; on the other hand, there is constructive mathematics with its intuitionistic logic and computational-content-by-design. In short, there are two opposing camps (classical and constructive) in mathematics separated by a no-man's land, with the occasional volley exchanged as in e.g. [6, 72, 90, 91].

Stimulated by Brouwer's revolutionary spirit, our goal is to subvert the above received view. To this end, we will identify a field of classical mathematics which occupies the twilight-zone between constructive and classical mathematics. Perhaps ironically, this very field is *Nonstandard Analysis*, so vilified by Bishop for its alleged *fundamentally* non-constructive nature. We introduce a well-known axiomatic approach to Nonstandard Analysis, Nelson's *internal set theory*, in Section 2. We discuss past claims regarding the constructive nature of Nonstandard Analysis in Section 3, including Osswald's notion of *local constructivity*.

In Section 4, we establish that Nonstandard Analysis occupies the twilight-zone between constructive and classical mathematics. As part of this endeavour, we show

[2]The naked noun *Nonstandard Analysis* will always implicitly include the adjective *classical*, i.e. based on classical logic. We shall not directly deal with *constructive Nonstandard Analysis*, i.e. based on intuitionistic logic, but do discuss its relationship with our results in Section 5.3.

[3]There are a number of approaches to constructive mathematics, as discussed at length in e.g. [7, III], [98, I.4], or [17], and both Brouwer's *intuitionism* and Bishop's *Constructive Analysis* ([11]) represent a school of constructive mathematics.

that a major part of *classical* Nonstandard Analysis has computational content, much like constructive mathematics itself. Intuitively, the predicate 'x is standard' unique to Nonstandard Analysis can be interpreted as 'x is computable', giving rise to computable (and even constructive) mathematics obtained directly from *classical* Nonstandard Analysis. Our results formalise Osswald's longstanding conjecture from Section 3.4 that classical Nonstandard Analysis is *locally constructive*. By way of reversal, we establish in Section 4.5 that certain *highly constructive* results extracted from Nonstandard Analysis in turn apply the nonstandard theorems from which they were obtained. Advanced results regarding the constructive nature of Nonstandard Analysis may be found in Section 5.

In Section 6, we shall offer an alternative interpretation of our results based on Brouwer's view that *logic is dependent on mathematics*. Indeed, Brouwer already explicitly stated in his dissertation that *logic depends upon mathematics* as follows:

> While thus mathematics is independent of logic, logic does depend upon mathematics: in the first place *intuitive logical reasoning* is that special kind of mathematical reasoning which remains if, considering mathematical structures, one restricts oneself to relations of whole and part; ([20, p. 127]; emphasis in Dutch original)

This leads us to the following alternative interpretation: If one fundamentally changes mathematics, as one arguably does when introducing Nonstandard Analysis, it stands to reason that the associated logic will change along as the latter depends on the former, in Brouwer's view. By way of an example, we shall observe in Section 4.1 that when introducing the notion of 'being standard' fundamental to Nonstandard Analysis, the law of excluded middle of classical mathematics moves from 'the original sin of non-constructivity' to a *computationally inert* principle which does not have any real non-constructive consequences anymore.

Thus, our results vindicate Brouwer's thesis that logic depends upon mathematics by showing that classical logic becomes 'much more constructive' when shifting from 'usual' mathematics to Nonstandard Analysis, as discussed in Section 6.1. As also discussed in the latter, Nelson and Robinson have indeed claimed that Nonstandard Analysis constitutes a new kind of mathematics and logic. We also discuss the connection between Nonstandard Analysis and Troelstra's views of intuitionism, and Brouwer's *first act* of intuitionism.

Finally, we point out a veritable *Catch22* connected to this paper. On one hand, if we were to emphasise that large parts of Nonstandard Analysis give rise to computable mathematics (and vice versa), then certain readers will be inclined to dismiss Nonstandard Analysis as 'nothing new', like Halmos was wont to:

> it's a special tool, too special, and other tools can do everything it does. It's all a matter of taste. ([39, p. 204])

On the other hand, if we were to emphasise the new mathematical objects (no longer involving Nonstandard Analysis) with strange properties one can obtain from Nonstandard Analysis, then certain readers will be inclined to dismiss Nonstandard Analysis as 'weird' or 'fundamentally different' from (mainstream) mathematics, as suggested by Bishop and Connes (See Section 3.5). To solve this conundrum, we shall walk a tightrope between new and known results, and hope all readers agree that neither of the aforementioned two views is correct.

2. An introduction to Nonstandard Analysis

We provide an informal introduction to Nonstandard Analysis in Section 2.1. Furthermore, we discuss Nelson's *internal set theory*, an axiomatic approach to Nonstandard Analysis, in detail in Section 2.2. Fragments of internal set theory based on *Peano* and *Heyting arithmetic* are studied in Section 2.3. These fragments, called H and P, are essential to our enterprise.

2.1. **An informal introduction to Nonstandard Analysis.**

In a nutshell, *Nonstandard Analysis* (NSA) is the formalisation of the intuitive *infinitesimal calculus*. The latter is used to date in large parts of physics and engineering, and was used historically in mathematics by e.g. Archimedes, Leibniz, Euler, and Newton until the advent of the Weierstraß 'epsilon-delta' framework, as detailed in [30].

Robinson first formulated the *semantic* approach to NSA around 1965 using *nonstandard models*, which are nowadays built using *free ultrafilters* ([74, 108]). Nelson later formulated a *syntactic* -or *axiomatic*- version of NSA in [62] called *internal set theory* (IST). In each approach, there is a universe of *standard* (usual/everyday) mathematical objects and a universe of *nonstandard* (new/other-worldly) mathematical objects, including infinitesimals. Tao has formulated a *pragmatic* view of NSA in [94, 95] which summarises as:

NSA is another tool for mathematics.

A sketch of Tao's view is as follows: given a *hard* problem in (usual/standard) mathematics, pushing this problem into the nonstandard universe of NSA generally converts it to a *different and easier* problem, a basic example being 'epsilon-delta' definitions. A *nonstandard* solution can then be formulated and pushed back into the universe of standard/usual mathematics.

This 'pushing back-and-forth' between the standard and nonstandard universe is enabled by *Transfer* and *Standard Part* (aka *Standardisation*). Robinson established the latter as properties of his nonstandard models, while Nelson adopts them as axioms of IST. The following figure illustrates Tao's view and the practice of NSA, as also discussed in Section 3.2 to 3.4. Note that *Transfer* is also used sometimes in tandem with *Standard Part*, and vice versa.

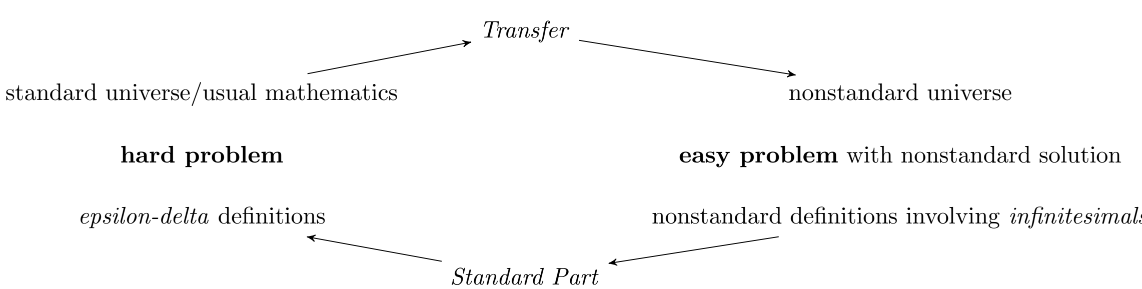

Figure 1. The practice of NSA

We shall study a number of basic examples of the scheme from Figure 1 in Section 4 for notions like continuity and Riemann integration. Furthermore, the following table lists some correspondences between the standard and nonstandard universe:

When pushed from the standard to the nonstandard universe, an object on the left corresponds to (or is included in) an object on the right side of the following table.

standard universe	**nonstandard universe**
continuous object	discrete object
infinite set	finite set
epsilon-delta definition	universal definition using infinitesimals
quantitative result	qualitative result

Finally, both *Transfer* and *Standard Part* are *highly non-constructive* in each approach to NSA, i.e. these principles imply the existence of *non-computable objects*. By contrast, the mathematics in the nonstandard universe is usually very basic, involving *little more than discrete sums and products* of nonstandard length. Various authors have made similar observations, which we discuss in Sections 3.2 to 3.4. In particular, Osswald has characterised this observation as **NSA is locally constructive** or the **local constructivity of NSA** as discussed in Section 3.4, and this view will be vindicated by our results in Section 4.

2.2. **An axiomatic approach to Nonstandard Analysis.** We shall introduce Nelson's *internal set theory* IST in this section.

2.2.1. *Internal Set Theory.* In Nelson's *syntactic* (or 'axiomatic') approach to Nonstandard Analysis ([62]), a new predicate 'st(x)', read as 'x is standard' is added to the language of ZFC, the usual foundation of mathematics[4]. The notations $(\forall^{\text{st}}x)$ and $(\exists^{\text{st}}y)$ are short for $(\forall x)(\text{st}(x) \rightarrow \dots)$ and $(\exists y)(\text{st}(y) \wedge \dots)$. A formula is called *internal* if it does not involve 'st', and *external* otherwise.

The external axioms *Idealisation*, *Standardisation*, and *Transfer* govern the new predicate 'st'; they are respectively defined[5] as follows:

- (I) $(\forall^{\text{st fin}}x)(\exists y)(\forall z \in x)\varphi(z, y) \rightarrow (\exists y)(\forall^{\text{st}}x)\varphi(x, y)$, for internal φ with any (possibly nonstandard) parameters.
- (S) $(\forall^{\text{st}}x)(\exists^{\text{st}}y)(\forall^{\text{st}}z)\big((z \in x \wedge \varphi(z)) \leftrightarrow z \in y\big)$, for any φ.
- (T) $(\forall^{\text{st}}t)\big[(\forall^{\text{st}}x)\varphi(x, t) \rightarrow (\forall x)\varphi(x, t)\big]$, where $\varphi(x, t)$ is internal, and only has free variables t, x.

The system IST is the *internal* system ZFC extended with the aforementioned *external* axioms; Internal set theory IST is a conservative extension of ZFC for the internal language ([62, §8]), i.e. these systems prove the same *internal* sentences.

For those familiar with Robinson's approach to Nonstandard Analysis, we point out one fundamental difference with internal set theory: In the former one studies *extensions* of structures provided by nonstandard models, while in the latter one studies the original structures, in which certain objects happen to be standard, as identified by 'st'. Nelson formulates this observation nicely as follows:

> *All theorems of conventional mathematics remain valid.* No change in terminology is required. What is new in internal set theory is only an addition, not a change. We choose to call certain sets standard (and we recall that in ZFC every mathematical object-a real

[4]The acronym ZFC stands for *Zermelo-Fraenkel set theory with the axiom of choice*.

[5]The superscript 'fin' in (I) means that x is finite, i.e. its number of elements are bounded by a natural number.

> number, a function, etc.-is a set), but the theorems of conventional mathematics apply to all sets, nonstandard as well as standard. ([62, p. 1165]; emphasis in original)
>
> *Every specific object of conventional mathematics is a standard set.* It remains unchanged in the new theory. For example, in internal set theory there is only one real number system, the system $\mathbb{R}$ with which we are already familiar. ([62, p. 1166]; emphasis in original)

Thus, Nelson's IST inherits non-constructive objects from ZFC, like the **standard** Turing jump[6], and axioms like the **internal** *law of excluded middle* $A \vee \neg A$.

It goes without saying that the step from ZFC to IST can be done for a large spectrum of logical systems weaker than ZFC. In Section 2.3, we study this extension of the usual classical and constructive axiomatisation of arithmetic, called *Peano Arithmetic* and *Heyting Arithmetic*. In the next sections, we discuss the intuitive meaning of the external axioms.

2.2.2. *Intuitive meaning of Idealisation.* Firstly, the *contraposition* of I implies that

$$(\forall y)(\exists^{\text{st}} x)\varphi(x,y) \to (\exists^{\text{st fin}} x)\underline{(\forall y)(\forall z \in x)\varphi(z,y)}, \tag{2.1}$$

for all internal φ, and where the underlined part is thus also *internal*. Hence, intuitively speaking, *Idealisation* allows us to 'pull a standard quantifiers like $(\exists^{\text{st}} x)$ in (2.1) through a normal quantifier $(\forall y)$'. We will also refer to (2.1) as *Idealisation*, as the latter axiom is usually used in this contra-posed form.

Note that the axioms $B\Sigma_n$ of Peano arithmetic ([23, II]) play a similar role: The former allow one to 'pull an unbounded number quantifier through a bounded number quantifier'. In each case, one obtains a formula in a kind of 'normal form' with a block of certain quantifiers (resp. external/unbounded) up front followed by another block of different quantifiers (resp. internal/bounded). Example 2.1 involving nonstandard continuity in IST is illustrative (See [62, §5] for more examples).

Example 2.1. We say that f is *nonstandard continuous* on the set $X \subseteq \mathbb{R}$ if

$$(\forall^{\text{st}} x \in X)(\forall y \in X)(x \approx y \to f(x) \approx f(y)), \tag{2.2}$$

where $z \approx w$ is $(\forall^{\text{st}} n \in \mathbb{N})(|z-w| < \frac{1}{n})$. Resolving '$\approx$' in (2.2), we obtain

$$(\forall^{\text{st}} x \in X)(\forall y \in X)\big((\forall^{\text{st}} N \in \mathbb{N})(|x-y| < \tfrac{1}{N}) \to (\forall^{\text{st}} k \in \mathbb{N})(|f(x)-f(y)| < \tfrac{1}{k})\big).$$

We may bring out the '$(\forall^{\text{st}} k \in \mathbb{N})$' and '$(\forall^{\text{st}} N \in \mathbb{N})$' quantifiers as follows:

$$(\forall^{\text{st}} x \in X)(\forall^{\text{st}} k \in \mathbb{N})\underline{(\forall y \in X)(\exists^{\text{st}} N \in \mathbb{N})\big(|x-y| < \tfrac{1}{N} \to |f(x)-f(y)| < \tfrac{1}{k}\big)}.$$

Applying I as in (2.1) to the underlined formula, we obtain a finite and standard set $z \subset \mathbb{N}$ such that $(\forall y \in X)(\exists N \in z)$ in the previous formula. Now let N_0 be the maximum of all numbers in z, and note that for $N = N_0$

$$(\forall^{\text{st}} x \in X)(\forall^{\text{st}} k \in \mathbb{N})(\exists^{\text{st}} N \in \mathbb{N})(\forall y \in X)\big(|x-y| < \tfrac{1}{N} \to |f(x)-f(y)| < \tfrac{1}{k}\big).$$

The previous formula has all standard quantifiers up front and is very close to the 'epsilon-delta' definition of continuity from mainstream mathematics. Hence, we observe the role of I: to connect the worlds of nonstandard mathematics (as in (2.2)) and mainstream mathematics.

[6]That the Turing jump is standard follows from applying *Transfer* to the internal statement 'there is a set of natural numbers which solves the Halting problem' formalised inside ZFC.

2.2.3. *Intuitive meaning of Transfer.* The axiom *Transfer* expresses that certain statements about *standard* objects are also true for *all* objects. This property is essential in proving the equivalence between epsilon-delta statements and their nonstandard formulation. The following example involving continuity is illustrative.

Example 2.2. Recall Example 2.1, the definition of nonstandard continuity (2.2) and the final equation in particular. Dropping the 'st' for N in the latter:

$$(\forall^{\text{st}} x \in X)(\forall^{\text{st}} k \in \mathbb{N})(\exists N \in \mathbb{N})(\forall y \in X)\big(|x - y| < \tfrac{1}{N} \to |f(x) - f(y)| < \tfrac{1}{k}\big).$$

Assuming X and f to be standard, we can apply T to the previous to obtain

$$(\forall x \in X)(\forall k \in \mathbb{N})(\exists N \in \mathbb{N})\underline{(\forall y \in X)|x - y| < \tfrac{1}{N} \to |f(x) - f(y)| < \tfrac{1}{k}} \tag{2.3}$$

Note that (2.3) is just the *usual epsilon-delta definition of continuity.* In turn, to prove that (2.3) implies nonstandard continuity as in (2.2), fix *standard* X, f, k in (2.3) and apply the contraposition of T to '$(\exists N \in \mathbb{N})\varphi(N)$' where φ is the underlined formula in (2.3). The resulting formula $(\exists^{\text{st}} N \in \mathbb{N})\varphi(N)$ immediately implies nonstandard continuity as in (2.2).

By the previous example, nonstandard continuity (2.2) and epsilon-delta continuity (2.3) are equivalent for standard functions in IST. However, the former involves far less quantifier alternations and is close to the intuitive understanding of continuity as 'no jumps in the graph of the function'. Hence, we observe the role of T: to connect the worlds of nonstandard mathematics (as in (2.2)) and mainstream mathematics (as in (2.3)), and we also have our first example of the practice of NSA from Figure 1 involving nonstandard and epsilon-delta definitions.

2.2.4. *Intuitive meaning of Standardisation.* The axiom *Standardisation* (also called *Standard Part*) is useful as follows: It is in general easy to build *nonstandard and approximate* solutions to mathematical problems in IST, but a *standard* solution is needed as the latter also exists in 'normal' mathematics (as it is suitable for *Transfer*). Intuitively, the axiom S tells us that from a *nonstandard approximate solution*, we can *always* find a *standard one.* Since we may (only) apply *Transfer* to formulas involving the latter, we can then also prove the latter is an object of normal mathematics. The following example is highly illustrative.

Example 2.3. The *intermediate value theorem* states that for every continuous function $f : [0, 1] \to [0, 1]$ such that $f(0)f(1) < 0$, there is $x \in [0, 1]$ such that $f(x) = 0$. Assuming f is standard, it is easy[7] to find a nonstandard real y in the unit interval such that $f(y) \approx 0$, i.e. y is an intermediate value 'up to infinitesimals'. The axiom S then tells[8] us that there is a *standard* real x such that $x \approx y$, and by the nonstandard continuity of f (See previous example), we have $f(x) \approx 0$. Now apply T to the latter[9] to obtain $f(x) = 0$. Hence, we have obtain the (internal) intermediate value theorem for standard functions, and T yields the full theorem.

[7]Since $[0, 1]$ is compact, we may assume that f is uniformly continuous there. Similar to Example 2.2, we may assume f is nonstandard uniformly continuous as in $(\forall x, y \in [0, 1])[x \approx y \to f(x) \approx f(y)]$. Let N be a nonstandard natural number and let $j \leq N$ be the least number such that $f(\frac{j}{N})f(\frac{j+1}{N}) \leq 0$. Then $f(j/N) \approx 0$ by nonstandard uniform continuity, and we are done.

[8]Let $y \in [0, 1]$ be such that $f(y) \approx 0$ and consider the set of rationals $z = \{q_1, q_1, q_2, \ldots, q_N\}$ where q_i is a rational such that $|y - q_i| < \frac{1}{i}$ and $q_i = \frac{j}{2^i}$ for some $j \leq i$, and N is a nonstandard number. Applying S, there is a standard set w such that $(\forall^{\text{st}} q)(q \in w \leftrightarrow q \in z)$. The *standard* sequence provided by w thus converges to a standard real $x \approx y$.

[9]Recall that '$f(x) \approx 0$' is short for $(\forall^{\text{st}} k \in \mathbb{N})(|f(x)| < \frac{1}{k})$.

Hence, we observe the role of S: to connect the worlds of nonstandard mathematics and mainstream (standard) mathematics by providing standard objects 'close to' nonstandard ones.

In conclusion, the external axioms of IST provide a connection between nonstandard and mainstream mathematics: They allow one to 'jump back and forth' between the standard and nonstandard world as sketched in Figure 1. This technique is useful as some problems (like switching limits and integrals) may be easier to solve in the discrete/finite world of nonstandard mathematics than in the continuous/infinite world of standard mathematics (or vice versa). This observation lies at the heart of Nonstandard Analysis and is a first step towards understanding its power. Nelson formulated this observation nicely as follows:

> *Part of the power of nonstandard analysis is due to the fact that a complicated internal notion is frequently equivalent, on the standard sets, to a simple external notion.* ([62, p. 1169])

2.3. **Fragments of Internal Set Theory.** Fragments of IST have been studied before and we are interested in the systems P and H introduced in [9].

In a nutshell, P and H are versions of IST based on the usual classical and intuitionistic axiomatisations of arithmetic, namely *Peano and Heyting arithmetic*. We refer to [54] for the exact definitions of our version of Peano and Heyting arithmetic, commonly abbreviated respectively as $\mathsf{E\text{-}PA}^\omega$ and $\mathsf{E\text{-}HA}^\omega$. In particular, the systems P and H are conservative[10] extensions of *Peano arithmetic* $\mathsf{E\text{-}PA}^\omega$ and *Heyting arithmetic* $\mathsf{E\text{-}HA}^\omega$, as also follows from Theorem 4.1. We discuss the systems P and H in detail in Sections 2.3.1 and 2.3.2, and list the axioms in full detail in Section A. We discuss why P and H are important to our enterprise in Section 4.1. Notations and definitions in P and H of common mathematical notions like natural and real numbers may be found in Section A.4.

2.3.1. *The classical system* P. We discuss the fragment P of IST from [9]. Similar to the way IST is an extension of ZFC, P is just the internal system $\mathsf{E\text{-}PA}^\omega$ with the language extended with a new standardness predicate 'st' and with some special cases of the external axioms of IST. The technical details of this extension may be found in Section A.2 while we now provide an intuitive motivation for the external axioms of P, assuming basic familiarity with the finite type system of Gödel's system T, also discussed in Section A.1.

First of all, the system P does not include any fragment of *Transfer*. The motivation for this omission is as follows: The system ACA_0 proves the existence of the non-computable *Turing jump* ([87, III]) and very weak fragments of T already imply versions of ACA_0. In particular, the following axiom is the *Transfer* axiom limited to universal number quantifiers:

$$(\forall^{\mathrm{st}} f^1)\big[(\forall^{\mathrm{st}} n^0) f(n) \neq 0 \to (\forall m) f(m) \neq 0\big]. \tag{Π^0_1-TRANS}$$

As proved in [81, §4.1] and Section 4.3, Π^0_1-TRANS is essentially the nonstandard version of the Turing jump. Hence, no *Transfer* is present in P as this axiom is fundamentally non-constructive, and would result in a non-conservative extension.

[10] Like for ZFC and IST, if the system P (resp. H) proves an internal sentence, then this sentence is provable in $\mathsf{E\text{-}PA}^\omega$ (resp. $\mathsf{E\text{-}HA}^\omega$).

Secondly, the system P involves the full axiom *Idealisation* in the language of P as follows: For any internal formula in the language of P:

$$(\forall^{\text{st}} x^{\sigma^*})(\exists y^\tau)(\forall z^\sigma \in x)\varphi(z,y) \to (\exists y^\tau)(\forall^{\text{st}} x^\sigma)\varphi(x,y), \tag{2.4}$$

As it turns out, the axiom I does not yield any 'non-constructive' consequences, which also follows from Theorem 4.1 below. As for (2.1), we will also refer to the contraposition of (2.4) as *Idealisation*, as the latter is used mostly in the contraposed form similar to (2.1) (but with finite sequences rather than finite sets).

Thirdly, the system P involves a weakening of *Standardisation*. In particular, the axiom S may be equivalently formulated as follows:

$$(\forall^{\text{st}} x)(\exists^{\text{st}} y)\Phi(x,y) \to \big(\exists^{\text{st}} F\big)(\forall^{\text{st}} x)\Phi(x, F(x)), \tag{S'}$$

for any formula Φ in the language IST. In this way, S may be viewed as a 'standard' version of the axiom of choice. In light of the possible non-constructive content of the latter (See e.g. [28]), it is no surprise that S has to be weakened. In particular, P includes the following version of S, called the *Herbrandised Axiom of Choice*:

$$(\forall^{\text{st}} x^\rho)(\exists^{\text{st}} y^\tau)\varphi(x,y) \to \big(\exists^{\text{st}} G^{\rho\to\tau^*}\big)(\forall^{\text{st}} x^\rho)(\exists y^\tau \in G(x))\varphi(x,y), \tag{$\mathsf{HAC}_{\mathsf{int}}$}$$

where φ is any internal formula in the language of P. Note that G does not output a witness for y, but a *finite list* of potential witnesses to y. This is quite similar to *Herbrand's theorem* ([23, I.2.5]), hence the name of $\mathsf{HAC}_{\mathsf{int}}$.

Finally, we list two basic but important axioms of P from Definition A.3 below. These axioms are inspired by Nelson's claim about IST as follows:

> *Every specific object of conventional mathematics is a standard set.*
> It remains unchanged in the new theory [IST]. ([62, p. 1166])

Specific objects of the system P obviously include the constants $0, 1, \times, +$, and anything built from those. Thus, the system P includes the following three axioms; we refer to Definition A.3 for the exact technical details.

(i) All constants in the language of $\mathsf{E\text{-}PA}^\omega$ are standard.
(ii) A standard functional applied to a standard input yields a standard output.
(iii) If two objects are equal and one is standard, so is the other one.

As a result, the system P proves that any term of $\mathsf{E\text{-}PA}^\omega$ is standard, which will turn out to be essential in Section 4.5. Note that 'equality' as in item (iii) only applies to 'actual' equality *and not equality on the reals* in P, as discussed in Remark A.12.

2.3.2. *The constructive system* H. We discuss the fragment H of IST from [9]. Similar to the way IST is an extension of ZFC, H is just the internal system $\mathsf{E\text{-}HA}^\omega$ with the language extended with a new standardness predicate 'st' and with some special cases of the external and internal axioms of IST.

The technical details of this extension may be found in Section A.3; we now provide an intuitive motivation for the external axioms of H, assuming basic familiarity with the finite type system of Gödel's system T (See Section A.1). Note that $\mathsf{E\text{-}HA}^\omega$ and H are based on *intuitionistic logic.*

First of all, the system H does not involve *Transfer* for the same reasons P does not. By contrast, the axioms $\mathsf{HAC}_{\mathsf{int}}$ and *Idealisation* as in (2.4) (and its contraposition) are included in H, with the restrictions on φ lifted even.

Secondly, the system H includes some 'non-constructive' axioms relativized to 'st'. We just mention the names of these axioms and refer to Section A.3 for a full description. The system H involves nonstandard versions of the following axioms: *Markov's principle* (See e.g. [7, p. 47]) and the *independence of premises* principle (See e.g. [54, §5]). *Nonetheless*, H proves the same internal sentence as $\mathsf{E}\text{-}\mathsf{HA}^{\omega}$ by Theorem 4.1, i.e. the nonstandard versions are not really non-constructive.

Finally, H also includes the basic axioms from Definition A.3 as listed at the end of Section 2.3.1. In particular, H proves that any term of $\mathsf{E}\text{-}\mathsf{HA}^{\omega}$ is standard.

3. The (non-)constructive nature of Nonstandard Analysis

3.1. **Introduction.** A number of informal claims have over the years been made about the constructive nature of Nonstandard Analysis. These claims range from short quotes, three of which listed below, to more detailed conjectures. Regarding the latter, we discuss in Sections 3.2 to 3.4 observations by Keisler, Wattenberg, and Osswald on the constructive nature of Nonstandard Analysis. Note that 'constructive' is often used as a synonym of 'effective' in this context.

> It has often been held that nonstandard analysis is highly non-constructive, thus somewhat suspect, depending as it does upon the ultrapower construction to produce a model [...] On the other hand, nonstandard *praxis* is remarkably constructive; having the extended number set we can proceed with explicit calculations. (Emphasis in original: [2, p. 31])

> On the other hand, nonstandard arguments often have in practice a very constructive flavor, much more than do the corresponding standard proofs. Indeed, those who use nonstandard arguments often say of their proofs that they are "constructive modulo an ultrafilter." This is especially true in measure theory. ([76, p. 230])

> The aim of this section is to describe a nonstandard map "explicitly". There is only one inconstructive step in the proof, namely the choice of a so-called ultrafilter. [...] Moreover, the fact that the approach is "almost constructive" has the advantage that in special cases one can better see what happens: Up to the ultrafilter one can "calculate" the nonstandard embedding $*$. ([103, p. 44])

By contrast, Connes and Bishop have claimed that Nonstandard Analysis is somehow *fundamentally non-constructive*; their views are discussed briefly in Section 3.5 for the sake of completeness, while a detailed study may be found in [82].

Finally, the attentive reader has noted that we choose not to discuss the claims of the 'French school' of Nonstandard Analysis, in particular Reeb and Harthong's *Intuitionnisme 84* ([40]), regarding the constructive nature of Nonstandard Analysis. We motivate this choice by the observation that the associated literature is difficult to access (even in the digital age) and is written in large part in rather academic French, which (no pun intended) has long ceased to be the *lingua franca*.

3.2. **Keisler's lifting method.** We discuss Keisler *lifting method* which is a template for the practice of Nonstandard Analysis involving the hyperreal line, and a special case of the scheme in Figure 1. Keisler formulates the lifting method as:

> The following strategy, sometimes called the *lifting method*, has been used to prove results which are formulated on the ordinary real line.
>
> Step 1 Lift the given 'real' objects up to internal approximations on the hyperfinite grid.
>
> Step 2 Make a series of hyperfinite computations to construct some new internal object on the hyperfinite grid.
>
> Step 3 Come back down to the real line by taking standard parts of the results of the computations.
>
> The hyperfinite computations in Step 2 will typically replace more problematic infinite computations on the real line. ([52, p. 234])

Keisler's description of the lifting method involves the words *constructive* and *computations*, which are elaborated upon as follows by Keisler:

> [...] hyperreal proofs seem to be more 'constructive' than classical proofs. For example, the solution of a stochastic differential equation given by the hyperreal proof is obtained by solving a hyperfinite difference equation by a simple induction and taking the standard part. The hyperreal proof is not constructive in the usual sense, because in ZFC the axiom of choice is needed even to get the existence of a hyperreal line. What often happens is that a proof within RZ or IST of a statement of the form $(\exists x)\phi(x)$ will produce an x which is definable from H, where H is an arbitrary infinite hypernatural number. Thus instead of a pure existence proof, one obtains an explicit solution except for the dependence on H. The extra information one gets from the explicit construction of the solution from H makes the proof easier to understand and may lead to additional results. ([52, p. 235])

We conclude that Keisler's lifting method reflects the ideas behind the scheme in Figure 1. Furthermore, we stress that one can define the notion of (non-)standard real number in the systems P and H from Section 2.3 and develop Nonstandard Analysis there including the Stone-Weierstraß theorem as in [81, §3-4]. In particular, most of ZFC is not needed to do (large parts of) Nonstandard Analysis, though one could (wrongly so) get the opposite impression from Keisler's previous quote.

3.3. **Wattenberg's algorithms.** Almost two decades ago, Wattenberg published the paper *Nonstandard Analysis and Constructivism?* in which he speculates on a connection between Nonstandard Analysis and constructive mathematics ([105]).

> This is a speculative paper. For some time the author has been struck by an apparent affinity between two rather unlikely areas of mathematics - nonstandard analysis and constructivism. [...] The purpose of this paper is to investigate these ideas by examining several examples. ([105, p. 303])

On one hand, with only slight modification, some of Wattenberg's theorems in Nonstandard Analysis are seen to yield effective and constructive theorems using the systems P and H from Section 2.3. On the other hand, some of Wattenberg's (explicit and implicit) claims regarding the constructive status of the axioms *Transfer* and *Standard Part* from Nonstandard Analysis can be shown to be incorrect. These results will be explored in [80], but we can make preliminary observations.

Wattenberg digs deeper than the quotes in Section 3.1 by making the following important observation regarding the praxis of Nonstandard Analysis.

> Despite an essential nonconstructive kernel, many nonstandard arguments are constructive until the final step, a step that frequently involves the standard part map. ([105, p. 303])

Wattenberg's observation echoes the scheme in Figure 1 from Section 2.1, namely that the axiom *Standardisation* (aka *Standard Part*) is used to push nonstandard solutions down to the world of standard/usual mathematics at the end of a nonstandard proof. However, in either approach to Nonstandard Analysis this axiom is highly non-constructive, and even very weak instances are such: Consider *Standardisation* as in (S') from Section 2.3.1 weakened as follows:

$$(\forall f^1)\big[(\forall^{\rm st}x^0)(\exists^{\rm st}y^0)(f(x)=_0 y)\rightarrow \big(\exists^{\rm st}F\big)(\forall^{\rm st}x)(f(x)=_0 F(x))\big]. \qquad (\mathsf{S}'')$$

It is proved in Section A.5 that $\mathsf{P}+\mathsf{S}''$ proves the non-constructive *weak König's lemma*. Furthermore, (S'') is translated to the *special fan functional* (First introduced in [78]; see Section 4.4) during term extraction, an object which is extremely hard[11] to compute, as studied in [65].

In a nutshell, Wattenberg suggests that the mathematics in the nonstandard universe in Figure 1 from Section 2.1 is rich in constructive content, but that stepping from the nonstandard to the standard universe using *Standardisation* is fundamentally non-constructive. Thus, the latter axiom should be avoided if one is interested in obtaining the constructive content of Nonstandard Analysis. Note however that Wattenberg (explicitly and implicitly) makes use of *Transfer* in [105], which is *also* highly non-constructive, as established in Section 4.3 and [81, §4]. In other words, the picture Wattenberg tries to paint is somewhat imperfect, which shall be remedied in the next section.

3.4. **Osswald's local constructivity.** We discuss the notion *local constructivity*, originally formulated by Osswald (See e.g. [108, §7], [67, §1-2], or [68, §17.5]). The following heuristic principle describes the notion of *local constructivity.*

Principle 3.1. *A mathematical proof is* ***locally constructive*** *if the core, the essential part, of the proof is constructive in nature.*

Intuitively, a proof P of a theorem T is *locally constructive* if we can omit a small number of non-constructive initial and/or final steps in P and obtain a proof Q of a constructive theorem T' very similar to T. In other words, a locally constructive proof has the form $P=(P',P'',P''')$, where P',P''' are inessential and may be non-constructive, and the 'main part' P'' is constructive. As it happens, isolating the constructive part P'' can lead to new results, as in the paper[12] [67] and Section 4.

[11] In particular, the special fan functional Θ from [78] is not computable (in the sense of Kleene's S1-S9) from the Suslin functional, the functional version of $\Pi^1_1\text{-}\mathsf{CA}_0$, as proved in [65].

[12] As a tentative example, the anonymous referee of [67] wrote in his report that techniques in the paper produce 'results, substantially extending those in the literature'.

Osswald has conjectured the **local constructivity** of (proofs in) classical **Nonstandard Analysis**. This conjecture is based on the following three observations regarding the praxis of Nonstandard Analysis:

(1) The first step in a proof in Nonstandard Analysis often involves *Transfer* to enter the nonstandard universe, e.g. to convert epsilon-delta definitions to nonstandard ones, like for epsilon-delta continuity as in Example 2.2. However, *Transfer* is non-constructive as established in Section 4.3.
(2) The mathematical practice of Nonstandard Analysis **after entering the nonstandard world** in large part consists of the manipulation of (nonstandard) universal formulas not involving existential quantifiers. This manipulation amounts to mere computation, in many cases nothing more than hyperdiscrete sums and products.
(3) The final step in a proof of Nonstandard Analysis often involves *Standardisation* to push nonstandard objects back into the standard universe, as also observed by Keisler and Wattenberg in the previous sections. However, this axiom is also non-constructive.

These three observations suggest the following: If we take a proof $P = (P', P'', P''')$ in Nonstandard Analysis where P' (resp. P''') collects the initial (resp. final) applications of *Transfer* (resp. of *Standardisation*), then P'' is a proof rich in constructive content. The proof P is thus **locally constructive** in the sense that P'' contains the core argument, and is constructive in a certain as-yet-undefined sense (which will be formalised in Section 4). Furthermore, to avoid the use of *Transfer*, we should avoid epsilon-definitions and work instead directly with the nonstandard definitions involving infinitesimals, as also discussed in Section 5.3.

In conclusion, Osswald has conjectured that **Nonstandard Analysis is locally constructive**, suggesting that what remains of Nonstandard Analysis after stripping away *Transfer* and *Standardisation* is constructive in some sense. We will formalise this observation in Section 4 by introducing *pure* Nonstandard Analysis in Definition 4.27.

3.5. **The Bishop-Connes critique of Nonstandard Analysis.** For completeness, we briefly discuss the critique of Nonstandard Analysis by Bishop and Connes, relating to the Bishop quote in Section 1. Their critique may be summarised as:

The presence of ideal objects (*in particular infinitesimals*) *in Nonstandard Analysis yields the absence of computational content.*

In particular, for rather different reasons and in different contexts, Bishop and Connes equate 'meaningful mathematics' and 'mathematics with computational content', and therefore claim Nonstandard Analysis is devoid of meaning as it lacks -in their view- any and all computational content. A detailed study of this topic by the author, debunking the Bishop-Connes critique, may be found in [82].

The often harsh criticism of Nonstandard Analysis by Bishop and Connes is a matter of the historical record and discussed in remarkable detail in [3, 26, 27, 48–50, 93]. Furthermore, the *arguments by Connes and Bishop for* this critique have been dissected in surprising detail *and found wanting* in the aforementioned references, but establishing the *opposite* of the Bishop-Connes critique, namely that Nonstandard Analysis is *rich in computational content*, first took place in [81, 82].

As noted in Section 1, we are not interested in adding to this literature, but we shall examine the motivations of Connes and Bishop for their critique, as this will provide more insight into Nonstandard Analysis. To this end, recall that Bishop considers his mathematics to be concerned with *constructive objects*, i.e. those described by algorithms on the integers. The following quote is telltale.

> Everything attaches itself to number, and every mathematical statement ultimately expresses the fact that if we perform certain computations within the set of positive integers, we shall get certain results. ([11, §1.1])

This quote suggest a *fundamental* ontological divide between Bishop's mathematics and Nonstandard Analysis, as the latter *by design* is based on ideal objects (like infinitesimals) with prima facia *no algorithmic description at all*. In this light, it is not a stretch of the imagination to classify Nonstandard Analysis as *fundamentally non-constructive*, i.e. antipodal to Bishop's mathematics, or to paraphrase Bishop: *an attempt to dilute meaning/computational content further*.

Furthermore, this 'non-constructive first impression' is confirmed by the fact that the 'usual' development of Nonstandard Analysis involves *quite non-constructive*[13] axioms. The latter is formulated by Katz and Katz as:

> [...] the hyperreal approach incorporates an element of non-constructivity at the basic level of the very number system itself. ([50, §3.3])

Finally, the aforementioned 'first impression' does not seem to disappear if one considers more basic mathematics, e.g. arithmetic rather than set theory. Indeed, *Tennenbaum's theorem* [51, §11.3] 'literally' states that any nonstandard model of Peano Arithmetic is not computable. *What is meant* is that for a nonstandard model $\mathcal{M}$ of Peano Arithmetic, the operations $+_{\mathcal{M}}$ and $\times_{\mathcal{M}}$ cannot be computably defined in terms of the operations $+_{\mathbb{N}}$ and $\times_{\mathbb{N}}$ of the standard model $\mathbb{N}$ of Peano Arithmetic. In light of the above, Nonstandard Analysis *seems* fundamentally non-constructive even at the level of arithmetic.

Now, the Fields medallist Alain Connes has formulated similar negative criticism of classical Nonstandard Analysis in print on at least seven occasions. The first table in [48, §3.1] runs a tally for the period 1995-2007. Connes judgements range from *inadequate* and *disappointing*, to *a chimera* and *irremediable defect*. Regarding the effective content of Nonstandard Analysis, the following quote is also telltale.

> Thus a non-standard number gives us canonically a non-measurable subset of $[0, 1]$. This is the end of the rope for being 'explicit' since (from another side of logics) one knows that it is just impossible to construct explicitely a non-measurable subset of $[0, 1]$! (Verbatim copy of the text in [24])

In contrast to Bishop, Connes does not take a foundational position but has more pragmatic arguments in mind: According to Connes, Nonstandard Analysis is fundamentally non-constructive and thus useless for physics, as follows:

> The point is that as soon as you have a non-standard number, you get a non-measurable set. And in Choquet's circle, having well

[13]The usual development of Robinson's Nonstandard Analysis proceeds via the construction of a nonstandard model using a free ultrafilter. The existence of the latter is only slightly weaker than the axiom of choice of ZFC ([108]).

> studied the Polish school, we knew that every set you can name is measurable. So it seemed utterly doomed to failure to try to use non-standard analysis to do physics. ([25, p. 26])

Indeed, a major aspect of physics is the testing of hypotheses against experimental data, nowadays done on computers. But how can this testing be done if the mathematical formalism at hand is fundamentally non-constructive, as Connes claims?

In Section 4.1, we establish the opposite of the Bishop-Connes critique: rather than being 'fundamentally non-constructive', infinitesimals and other nonstandard objects provide an *elegant shorthand* for computational/constructive content.

4. Jenseits von Konstruktiv und Klassisch

In this section, we shall establish that Nonstandard Analysis occupies the twilight zone between the constructive and non-constructive. Perhaps more accurately, as also suggested by the title[14] of this section, Nonstandard Analysis will be shown to *transcend* the latter distinction.

To this end, we explore the constructive status of H extended with classical logic in Section 4.1 and shall discover that these extensions can neither rightfully be called non-constructive or constructive. What is more, the observations from Section 4.1 pertaining to logic allow us to uncover the vast computational content of classical Nonstandard Analysis. We shall consider three examples in Section 4.2 to 4.4 and formulate a general template in Section 4.6 for obtaining computation from *pure* Nonstandard Analysis, as defined in Definition 4.27. These results will establish the **local constructivity** of Nonstandard Analysis as formulated in Section 3.4. As a crowning achievement, we show in Section 4.5 that theorems of Nonstandard Analysis are 'meta-equivalent' to theorems rich in computational content.

4.1. **The twilight zone of mathematics.** In this section, we study some computational and constructive aspects of the systems H and P. During this study, it will turn out that the latter occupies the twilight zone between the constructive and non-constructive, as suggested in Section 1.

We first consider the system H, the conservative extension of the constructive system $\mathsf{E\text{-}HA}^\omega$ from the previous section. Now, if LEM consists of $A \vee \neg A$ for all internal A, then $\mathsf{E\text{-}HA}^\omega + \mathsf{LEM}$ is just $\mathsf{E\text{-}PA}^\omega$. In other words, the addition of LEM turns $\mathsf{E\text{-}HA}^\omega$ into a *classical* system, losing the constructive (BHK) interpretation of the quantifiers. Now consider $\mathsf{H} + \mathsf{LEM}$ and recall the following theorem, which is called the *main theorem on program extraction* in [9, Theorem 5.9].

Theorem 4.1 (Term extraction I)**.** *Let φ be internal, i.e. not involving* 'st'*, and let Δ_{int} be a collection of internal formulas. If $\mathsf{H} + \Delta_{\mathsf{int}} \vdash (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\varphi(x, y)$, then we can extract a term t from this proof with $\mathsf{E\text{-}HA}^\omega + \Delta_{\mathsf{int}} \vdash (\forall x)(\exists y \in t(x))\varphi(x, y)$. The term t is such that $t(x)$ is a finite list and does not depend on Δ_{int}.*

Note that the *conclusion* of the theorem, namely '$\mathsf{E\text{-}HA}^\omega + \Delta_{\mathsf{int}} \vdash (\forall x)(\exists y \in t(x))\varphi(x, y)$', *does not involve Nonstandard Analysis* anymore. The term t from the previous theorem is part of Gödel's system T from Section A.1, i.e. essentially a computer program formulated in e.g. Martin-Löf type theory or Agda ([1, 60]).

[14] We leave it to the reader to decide how our title corresponds to Nietzsche's famous book.

What is more important is the following observation (stemming from the proof of [9, Theorem 5.5]): Taking Δ_{int} to be empty, we observe that constructive information (the term t) can be obtained from proofs in H. This is a nice result, but hardly surprising in light of the existing semantic approach to *constructive* Nonstandard Analysis (See Section 5.3) and the realisability interpretation of constructive mathematics. What is surprising is the following: Taking Δ_{int} to be LEM in Theorem 4.1, we observe that proofs in $\mathsf{H} + \mathsf{LEM}$ or H provide *the same kind of constructive information*, namely the term t. Of course, the final system (namely $\mathsf{E\text{-}HA}^{\omega} + \mathsf{LEM}$) is non-constructive in the former case, but the fact remains that the *original sin of non-constructivity* LEM *does not influence* the constructive information (in the form of the term t) extracted from H, as is clear from Theorem 4.1.

Hence, while $\mathsf{E\text{-}HA}^{\omega}$ looses its constructive nature when adding LEM, the system H retains its 'term extraction property' by Theorem 4.1 when adding LEM. To be absolutely clear, we do **not** claim that $\mathsf{H}+\mathsf{LEM}$ is a constructive system, but observe that it has the *same term extraction property* as the constructive system H. For this reason, $\mathsf{H}+\mathsf{LEM}$ is neither constructive (due to the presence LEM) nor can it be called non-constructive (due to the presence of the **same** term extraction property as the *constructive* system H). In conclusion, $\mathsf{H} + \mathsf{LEM}$ is an example of a system which inhabits the *twilight zone* between the constructive and non-constructive.

An alternative view based on Brouwer's dictum *logic depends upon mathematics* is that extending $\mathsf{E\text{-}HA}^{\omega}$ to the nonstandard system H *fundamentally* changes the mathematics. The logic, dependent as it is on mathematics according to Brouwer, changes along too; an example of this change-in-logic is as follows: LEM changes from *the original sin of non-constructivity* in the context of $\mathsf{E\text{-}HA}^{\omega}$ to a *computationally inert* statement in the context of H by Theorem 4.1. We shall discuss this view in more detail in Section 6.1.

As a logical next step, we discuss P, which is essentially $\mathsf{H} + \mathsf{LEM}_{\mathsf{ns}}$ and where the latter is the law of excluded middle $\Phi \vee \neg\Phi$ for *any* formula (in the language of H). Note that $\mathsf{LEM}_{\mathsf{ns}}$ is much more general than LEM. *Nonetheless*, we have the following theorem, not proved explicitly in [9], and first formulated in [81].

Theorem 4.2 (Term extraction II)**.** *Let φ be internal, i.e. not involving* 'st'*, and let Δ_{int} be a collection of internal formulas. If $\mathsf{P} + \Delta_{\mathsf{int}} \vdash (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\varphi(x, y)$, then we can extract a term t from this proof with $\mathsf{E\text{-}PA}^{\omega} + \Delta_{\mathsf{int}} \vdash (\forall x)(\exists y \in t(x))\varphi(x, y)$. The term t is such that $t(x)$ is a finite list and does not depend on Δ_{int}.*

As for $\mathsf{H}+\mathsf{LEM}$, the system P is neither constructive due to the presence LEM *and* $\mathsf{LEM}_{\mathsf{ns}}$, but it has the **same** term extraction property as the constructive system H by Theorem 4.2, i.e. calling P non-constructive does not make sense in light of (and does not do justice to) Theorem 4.1 for $\Delta_{\mathsf{int}} = \varnothing$.

In short, the systems $\mathsf{H} + \mathsf{LEM}$ and P share essential properties with classical mathematics (in the form of LEM and $\mathsf{LEM}_{\mathsf{ns}}$) and with constructive mathematics (the term extraction property of H from Theorem 4.1). For this reason, these systems can be said to occupy the *twilight zone* between the constructive and non-constructive: They are neither one nor the other. More evidence for the latter claim may be found in Section 5.2.1 where P is shown to satisfy a natural nonstandard version of the *existence property* and related 'hallmark' properties of intuitionistic systems. We show in Section 4.3 that adding *Transfer* to P results in an unambiguously *non-constructive* system; the axiom *Transfer* will be shown to be the 'real' law of excluded middle of Nonstandard Analysis in a very concrete way.

Furthermore, on a more pragmatic note, the above results betray that we can gain access to the computational/constructive content of *classical* Nonstandard Analysis via the term extraction property formulated in Theorem 4.2. It is then a natural question *how wide the scope of the latter is*. The rest of this section is dedicated to showing that this scope encompasses *pure Nonstandard Analysis*. The latter is Nonstandard Analysis restricted to *nonstandard axioms* (*Transfer* and *Standardisation*) and *nonstandard definitions* (for continuity, compactness, integration, convergence, et cetera). This will establish the **local constructivity of Nonstandard Analysis** as formulated in Section 3.4.

The above discussion deals with *logic* and we now discuss a first *mathematical* example. In particular, we now consider an elementary application of Theorem 4.2 to *nonstandard continuity* as in Example 2.1. We shall refer to a formula of the form $(\forall^{\text{st}}x)(\exists^{\text{st}}y)\varphi(x,y)$ with φ internal as a *normal form*.

Example 4.3 (Nonstandard and constructive continuity)**.** Suppose $f:\mathbb{R}\to\mathbb{R}$ is nonstandard continuous, provable in P. In other words, similar to Example 2.1, the following is provable in our system P:

$$(\forall^{\text{st}}x\in\mathbb{R})(\forall y\in\mathbb{R})(x\approx y\to f(x)\approx f(y)). \tag{4.1}$$

Since P includes *Idealisation* (essentially) as in IST, P also proves the following:

$$(\forall^{\text{st}}x\in\mathbb{R})(\forall^{\text{st}}k\in\mathbb{N})(\exists^{\text{st}}N\in\mathbb{N}) \tag{4.2}$$
$$\underline{(\forall y\in\mathbb{R})\big(|x-y|<\tfrac{1}{N}\to|f(x)-f(y)|<\tfrac{1}{k}\big)},$$

in exactly the same way as proved in Example 2.1. Since the underlined formula in (4.2) is internal, (4.2) is a normal form and we note that Theorem 4.2 applies to '$\mathsf{P}\vdash(4.2)$'. Applying the latter, we obtain a term t such that $\mathsf{E\text{-}PA}^\omega$ proves:

$$(\forall x\in\mathbb{R},k\in\mathbb{N})(\exists N\in t(x,k))(\forall y\in\mathbb{R})\big(|x-y|<\tfrac{1}{N}\to|f(x)-f(y)|<\tfrac{1}{k}\big).$$

Since $t(x,k)$ is a finite list of natural numbers, define $s(x,k)$ as the maximum of $t(x,k)(i)$ for $i<|t(x,k)|$ where $|t(x,k)|$ is the length of the finite list $t(x,k)$. The term s is called a *modulus* of continuity of f as it satisfies:

$$(\forall x\in\mathbb{R},k\in\mathbb{N})(\forall y\in\mathbb{R})\big(|x-y|<\tfrac{1}{s(x,k)}\to|f(x)-f(y)|<\tfrac{1}{k}\big). \tag{4.3}$$

Similarly, from the proof in P that f is nonstandard *uniform* continuous, we may extract a modulus of *uniform* continuity (See Section 4.2). This observation is *important*: moduli are an essential part of Bishop's *Constructive Analysis* (See e.g. [11, p. 34]), so we just proved that such constructive information *is implicit in the nonstandard notion of continuity*!

For the 'reverse direction', the basic axioms (See Definition A.3) of P state that all constants in the language of $\mathsf{E\text{-}PA}^\omega$ are standard. Hence, if $\mathsf{E\text{-}PA}^\omega$ proves (4.3) for some term s, then P proves (4.3) **and** that s is standard. Hence, P proves

$$(\forall^{\text{st}}x\in\mathbb{R},k\in\mathbb{N})(\exists^{\text{st}}N\in\mathbb{N})(\forall y\in\mathbb{R})\big(|x-y|<\tfrac{1}{N}\to|f(x)-f(y)|<\tfrac{1}{k}\big), \tag{4.4}$$

as we may take $N=s(x,k)$. However, (4.4) clearly implies (4.1), i.e. nonstandard continuity also follows from 'constructive' continuity with a modulus! Furthermore, one establishes in exactly the same way that $\mathsf{P}\vdash[(4.3)\to(4.1)]$ for any term t.

By the previous example, the nonstandard notion of (uniform) continuity amounts to the same as the 'constructive' definition involving moduli. As discussed in Section 4.6, other nonstandard definitions (of integration, differentiability, compactness, convergence, et cetera) are also 'meta-equivalent' to their constructive counterparts. Thus, rather than being 'fundamentally non-constructive' as claimed by Bishop and Connes in Section 3.5, infinitesimals (and other nonstandard objects) provide an *elegant shorthand* for computational content, like moduli from constructive mathematics. Furthermore, the types involved in nonstandard statements are generally lower (than the associated statement not involving Nonstandard Analysis), which should appeal to the practitioners of *Reverse Mathematics* where higher-type objects are represented via so-called codes in second-order arithmetic.

Finally, Theorem 4.2 allows us to extract computational content (in the form of the term t) from (classical) Nonstandard Analysis in the spirit of Kohlenbach's *proof mining* program. We refer to [54] for the latter, while we shall apply Theorem 4.2 in the next sections to establish the local constructivity of Nonstandard Analysis.

4.2. **Two examples of local constructivity.** In this section, we study two basic examples of theorems proved using Nonstandard Analysis, namely the *intermediate value theorem* and *continuous functions on the unit interval are Riemann integrable*. We shall apply the heuristic provided by *local constructivity*, i.e. only consider the parts of the proof in the nonstandard universe not involving *Transfer* and *Standard Part*, and apply Theorem 4.2 to obtain the effective content.

4.2.1. *Intermediate value theorem.* In this section, we discuss the computational content of a nonstandard proof of the *intermediate value theorem*; we shall make use of the *uniform* notion of nonstandard continuity, defined as follows.

Definition 4.4. The function f is *nonstandard uniformly continuous* on $[0,1]$ if

$$(\forall x, y \in [0,1])[x \approx y \to f(x) \approx f(y)]. \tag{4.5}$$

We first prove the *intermediate value theorem* in IST; all notions have their usual (epsilon-delta/internal) meaning, unless explicitly stated otherwise.

Theorem 4.5 (IVT)**.** *For every continuous function $f : [0,1] \to [0,1]$ such that $f(0)f(1) < 0$, there is $x \in [0,1]$ such that $f(x) = 0$.*

Proof. Clearly IVT is internal, and we may thus assume that f is standard, as applying *Transfer* to this restriction yields full IVT. By Example 2.2, f is nonstandard continuous as in (2.2) for $X = [0,1]$ by *Transfer*. The latter can also be used to show that f is nonstandard *uniform* continuous as in (4.1), but this follows easiest from $(\forall x \in [0,1])(\exists^{\text{st}} y \in [0,1])(x \approx y)$, i.e. the nonstandard compactness of the unit interval, which follows in turn from applying *Standard Part* to the trivial formula $(\forall^{\text{st}} k \in \mathbb{N})(\exists^{\text{st}} q \in \mathbb{Q})(|x - q| < \frac{1}{2^k})$ and noting that the resulting sequence converges (to a standard real). Now let N be a nonstandard natural number and let $j \leq N$ be the least number such that $f(\frac{j}{N})f(\frac{j+1}{N}) \leq 0$ (whose existence is guaranteed by $f(0)f(1) < 0$ and induction). Then $f(j/N) \approx 0$ by nonstandard uniform continuity. By nonstandard compactness there is standard $y \in [0,1]$ such that $y \approx j/N$. Then $f(y) \approx 0$ by nonstandard uniform continuity, and the former is short for $(\forall^{\text{st}} k^0)(|f(x)| < \frac{1}{k})$. Finally, applying *Transfer* yields $f(x) = 0$. $\square$

Local constructivity as in Section 3.4 dictates that we omit all instances of *Transfer* and *Standard Part* to obtain the constructive core of the nonstandard proof of IVT. Note that the former axioms were used to obtain nonstandard *uniform* continuity. Thus, we obtain the following version, which we prove in P.

Theorem 4.6 ($\mathsf{IVT}_{\mathsf{ns}}$). *For every standard and nonstandard uniformly continuous $f : [0,1] \to [0,1]$ such that $f(0)f(1) < 0$, there is $x \in [0,1]$ such that $f(x) \approx 0$.*

Proof. Let f be as in the theorem, fix $N^0 > 0$ and suppose for all $i \leq N$ that $f(\frac{i}{N})f(\frac{i+1}{N}) > 0$. If $f(0) > 0$, let $\varphi(i)$ be the formula $i \leq N \to f(i/N) > 0$. Then $\varphi(0)$ and $\varphi(i) \to \varphi(i+1)$ for all $i \in \mathbb{N}$, both by assumption. Induction now yields that $f(1) > 0$, a contradiction. Thus, let N be a nonstandard natural number and let $j \leq N$ be the least number such that $f(\frac{j}{N})f(\frac{j+1}{N}) \leq 0$. Then $f(j/N) \approx 0$ by nonstandard uniform continuity. The same approach works for $f(0) < 0$. $\square$

Applying *term extraction* to $\mathsf{IVT}_{\mathsf{ns}}$ as in Theorem 4.2, we obtain the following theorem, which includes a *constructive* 'approximate' version of IVT (See [7, I.7]).

Theorem 4.7. *From the proof '$\mathsf{P} \vdash \mathsf{IVT}_{\mathsf{ns}}$', we obtain a term t such that $\mathsf{E\text{-}PA}^\omega$ proves that for all f, g where g is the modulus of uniform continuity of f, we have $(\forall k \in \mathbb{N})(|f(t(f,g,k))| <_{\mathbb{R}} \frac{1}{k})$.*

Proof. We may apply Theorem 4.2 to normal forms only. Now, the nonstandard uniform continuity of f on $[0,1]$ has the following normal form:

$$(\forall^{\text{st}} k \in \mathbb{N})(\exists^{\text{st}} N \in \mathbb{N})\underline{(\forall x, y \in [0,1])\big(|x-y| < \tfrac{1}{N} \to |f(x) - f(y)| < \tfrac{1}{k}\big)}, \tag{4.6}$$

which is proved in P exactly as for nonstandard continuity in Example 2.1; the underlined formula is abbreviated $A(f, N, k)$. The conclusion of $\mathsf{IVT}_{\mathsf{ns}}$ is not a normal form, but is *equivalent*[15] to $(\forall^{\text{st}} k \in \mathbb{N})(\exists x \in [0,1])(|f(x)| < \frac{1}{k})$ by *Idealisation*. By nonstandard uniform continuity, we also have $(\forall^{\text{st}} k \in \mathbb{N})(\exists^{\text{st}} q \in [0,1])(|f(q)| < \frac{1}{k})$, which *is* a normal form. Hence, taking '$f \in D$' to mean that $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(0)f(1) < 0$, $\mathsf{IVT}_{\mathsf{ns}}$ becomes

$$(\forall^{\text{st}} f \in D)\big[(\forall^{\text{st}} l^0)(\exists^{\text{st}} N^0)A(f, N, l) \to (\forall^{\text{st}} k^0)(\exists^{\text{st}} q \in [0,1])(|f(q)| < \tfrac{1}{k})\big]. \tag{4.7}$$

Since standard functions output standard values for standard inputs, we obtain

$$(\forall^{\text{st}} f \in D, g)\big[(\forall^{\text{st}} l^0)A(f, g(l), l) \to (\forall^{\text{st}} k^0)(\exists^{\text{st}} q \in [0,1])(|f(q)| < \tfrac{1}{k})\big], \tag{4.8}$$

and dropping the remaining 'st' in the antecedent, we get

$$(\forall^{\text{st}} f \in D, g)\big[(\forall l^0)A(f, g(l), l) \to (\forall^{\text{st}} k^0)(\exists^{\text{st}} q \in [0,1])(|f(q)| < \tfrac{1}{k})\big], \tag{4.9}$$

which becomes the following normal form:

$$(\forall^{\text{st}} f \in D, g, k)(\exists^{\text{st}} q \in [0,1])\big[(\forall l^0)A(f, g(l), l) \to (|f(q)| < \tfrac{1}{k})\big]. \tag{4.10}$$

Applying Theorem 4.2 to '$\mathsf{P} \vdash$ (4.10)', we obtain a term s such that

$$(\forall f \in D, g, k)(\exists q \in s(f,g,k))\big[(\forall l^0)A(f, g(l), l) \to (|f(q)| < \tfrac{1}{k})\big]. \tag{4.11}$$

To obtain the term t form the theorem, note that we can decide if $|f(q)| > \frac{1}{2^k}$ or $|f(q)| < \frac{1}{k}$ for any q, using the 'law of comparison' ([7, §1.3]) of constructive mathematics. Indeed, this law states that (under the BHK-interpretation):

$$(\forall x, y, z \in \mathbb{R})(x < y \to x < z \vee z < y). \tag{4.12}$$

[15] Apply *Idealisation* to $(\forall^{\text{st}} z^{0^*})(\exists x \in [0,1])(\forall k \in z)(|f(x)| < \frac{1}{k})$ to see this.

Thus, there is an effective procedure to decide which disjunct in (4.12) holds. Hence, define $t(f, g, k)$ as the first element $q \in s(f, g, 2^k)$ such that $|f(q)| < \frac{1}{k}$. We obtain

$$(\forall f \in D, g)\big[(\forall l^0)A(f, g(l), l) \to (\forall k^0)(|f(t(f, g, k))| < \tfrac{1}{k})\big], \tag{4.13}$$

from (4.11) by pushing the quantifier pertaining to k inside again. □

Approximate versions of IVT are of course well-known, but the previous is not intended to be new, but merely to provide an example of local constructivity as in Section 3.4: After stripping away *Transfer* and *Standard Part* in the proof of Theorem 4.5, we are left with a proof of $\mathsf{IVT}_{\mathsf{ns}}$ in P, which converts into a constructive version of IVT after term extraction by Theorem 4.7.

Furthermore, we point out the (apparently innocent) use of proof-by-contradiction in the proof of $\mathsf{IVT}_{\mathsf{ns}}$. Also, Theorem 4.7 is merely *one of many* possible results obtainable by term extraction. Indeed, while the conclusion of the former theorem has effective content, we can as well obtain *relative computability results* from the proof of IVT in Theorem 4.4, e.g. by not removing the use of *Transfer*. We shall explore this avenue further in Section 4.3.2.

Finally, it is also important to note that we wrote out the proof of Theorem 4.7 in full detail, while one readily skips from (4.7) to (4.13) with some practice. To the latter end, we conclude this section with the following remark on normal forms; we show that moduli (like in (4.10)) come about when converting an implication between two normal forms into an normal form in P (See Theorem 5.5 for H).

Remark 4.8 (Normal forms and implication)**.** As discussed in Remark 4.3, the nonstandard definition of continuity can be brought into a normal form (4.2), while the nonstandard definition of *uniform* continuity has the normal from (4.6). These observations are important as normal forms have computational content thanks to Theorem 4.2. The proof of Theorem 4.7 suggests that *normal forms are closed under implication*; indeed, the normal form (4.10) is derived from an *implication* between two normal forms. We now establish the general case as follows. Let φ, ψ be internal and consider the following implication between normal forms:

$$(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\varphi(x, y) \to (\forall^{\mathrm{st}} z)(\exists^{\mathrm{st}} w)\psi(z, w). \tag{4.14}$$

Since standard functionals have standard output for standard input, (4.14) implies

$$(\forall^{\mathrm{st}} \zeta)\big[(\forall^{\mathrm{st}} x)\varphi(x, \zeta(x)) \to (\forall^{\mathrm{st}} z)(\exists^{\mathrm{st}} w)\psi(z, w)\big]. \tag{4.15}$$

Bringing all standard quantifiers outside, we obtain the following normal form:

$$(\forall^{\mathrm{st}} \zeta, z)(\exists^{\mathrm{st}} w, x)\big[\varphi(x, \zeta(x)) \to \psi(z, w)\big], \tag{4.16}$$

as the formula in square brackets is internal. Now, (4.16) is equivalent to (4.15), but one usually (like in the proof of Theorem 4.7) weakens the latter as follows:

$$(\forall^{\mathrm{st}} \zeta, z)(\exists^{\mathrm{st}} w)\big[(\forall x)\varphi(x, \zeta(x)) \to \psi(z, w)\big], \tag{4.17}$$

as (4.17) is closer to the usual mathematical definitions. For instance, if the antecedent of (4.14) is (the normal form of) uniform continuity, then the antecedent of (4.17) is the constructive definition of uniform-continuity-with-a-modulus, while this is not the case for (4.16). We shall shorten the remaining proofs in this paper by just providing normal forms and jumping straight from (4.14) to (4.17) whenever possible.

4.2.2. *Riemann integration.* In this section, we discuss the computational content of a nonstandard proof that continuous functions are Riemann integrable; we shall make use of the usual definitions of Riemann integration as follows.

Definition 4.9. [Riemann Integration]

(1) A *partition* of $[0,1]$ is an increasing sequence $\pi = (0, t_0, x_1, t_1, \dots, x_{M-1}, t_{M-1}, 1)$. We write '$\pi \in P([0,1])$' to denote that π is such a partition.
(2) For $\pi \in P([0,1])$, $\|\pi\|$ is the *mesh*, i.e. the largest distance between two adjacent partition points x_i and x_{i+1}.
(3) For $\pi \in P([0,1])$ and $f : \mathbb{R} \to \mathbb{R}$, the real $S_\pi(f) := \sum_{i=0}^{M-1} f(t_i)(x_{i+1} - x_i)$ is the *Riemann sum* of f and π.
(4) A function f is *nonstandard integrable* on $[0,1]$ if

$$(\forall \pi, \pi' \in P([0,1]))\big[\|\pi\|, \|\pi'\| \approx 0 \to S_\pi(f) \approx S_{\pi'}(f)\big]. \tag{4.18}$$

(5) A function f is *integrable* on $[0,1]$ if

$$(\forall k^0)(\exists N^0)(\forall \pi, \rho \in P([0,1]))\big[\|\pi\|, \|\rho\| < \tfrac{1}{N} \to |S_\pi(f) - S_\rho(f)| < \tfrac{1}{k}\big]. \tag{4.19}$$

A modulus of (Riemann) integration ω^1 provides $N = \omega(k)$ as in (4.19).

Let RIE be the (internal) statement that a function is Riemann integrable on the unit interval if it is (pointwise) continuous there.

Theorem 4.10. *The system* IST *proves* RIE.

Proof. In a nutshell, the proof in [45, p. 57] goes through with minimal modification (from the Robinsonian framework). In the latter proof, *Transfer* and *Standardisation* (although the latter can be avoided) are used to derive the nonstandard definitions of *uniform* continuity and integration from the epsilon-delta definitions of pointwise continuity and Riemann integration. We have proved the equivalence between epsilon-delta and nonstandard continuity in Remark 2.2, and the case for Riemann integration is essentially identical. It thus remains to prove that nonstandard uniform continuity implies nonstandard integrability.

Thus, let $\pi' = (0, t'_0, x'_1, t'_1, \dots, x'_{M-1}, t'_{M-1}, 1)$ be $\pi = (0, t_0, x_1, t_1, \dots, x_{M-1}, t_{M-1}, 1)$ with all reals replaced[16] with their rational approximations of the form $\frac{i}{2^{M+1}}$. By nonstandard uniform continuity, we have $\max_{i \le M-1} |f(t'_i) - f(t_i)| = \zeta_0 \approx 0$, and define $x'_{i+1} - x'_i = x_{i+1} - x_i + \varepsilon_i$ where $0 \approx |\varepsilon_i| \le \frac{1}{2^{M+1}}$ by definition. Applying (4.6) for $k = 1$ and suitable standard $x \in [0,1]$, we see that $f(y) \le M$ for any $y \in [0,1]$. Thus $|\sum_{i=0}^{M-1} f(t'_i)\varepsilon_i| \le \frac{M^2}{2^{M+1}} \approx 0$ and consider the following:

$$\begin{aligned}
|S_\pi(f) - S_{\pi'}(f)| &= |\textstyle\sum_{i=0}^{M-1} f(t_i)(x_{i+1} - x_i) - \sum_{i=0}^{M-1} f(t'_i)(x'_{i+1} - x'_i)| \\
&= |\textstyle\sum_{i=0}^{M-1} (f(t'_i) - f(t_i)) \cdot (x_{i+1} - x_i) + \sum_{i=0}^{M-1} f(t'_i)\varepsilon_i| \\
&\le \textstyle\sum_{i=0}^{M-1} |f(t'_i) - f(t_i)| \cdot |x_{i+1} - x_i| + |\sum_{i=0}^{M-1} f(t'_i)\varepsilon_i| \\
&\lessapprox \textstyle\sum_{i=0}^{M-1} \zeta_0 \cdot |x_{i+1} - x_i| \\
&= \zeta_0 \textstyle\sum_{i=0}^{M-1} \cdot |x_{i+1} - x_i| = \zeta_0 \approx 0.
\end{aligned} \tag{4.20}$$

For two partitions $\pi, \rho \in P([0,1])$, consider the associated 'approximate' partitions π' and ρ'. Since the latter only consist of rationals, it is easy to define 'refinements'

[16] Note that this can be done effectively by using the 'law of comparison' ([7, §1.3]) as in (4.12).

π'', ρ'' of equal length and which contain all points with even index from both π' and ρ'; the points of odd index have to be repeated as necessary. We then have

$$S_\pi(f) \approx S_{\pi'}(f) \approx S_{\pi''}(f) \approx S_{\rho''}(f) \approx S_{\rho'}(f) \approx S_\rho(f),$$

where only the third '$\approx$' requires a proof, which is analogous to (4.20). $\square$

Local constructivity as in Section 3.4 dictates that we omit all instances of *Transfer* and *Standard Part* to obtain the constructive core of the nonstandard proof of RIE. Note that the former axioms were used to obtain nonstandard *uniform* continuity and let $\mathsf{RIE}_{\mathsf{ns}}$ be the statement that *every nonstandard uniformly continuous function f on the unit interval is nonstandard integrable.*

Let $\mathsf{RIE}_{\mathsf{ef}}(t)$ be the statement that $t(g)$ *is a modulus of integration if g is a modulus of uniform continuity of f on the unit interval.* The latter theorem may be called 'constructive' as we copied it from Bishop's constructive analysis (See [16, p. 53]). The nonstandard version apparently yields the constructive version.

Theorem 4.11. *From a proof of $\mathsf{RIE}_{\mathsf{ns}}$ in P, a term t can be extracted such that $\mathsf{E}\text{-}\mathsf{PA}^\omega$ proves $\mathsf{RIE}_{\mathsf{ef}}(t)$.*

Proof. A normal form for nonstandard uniform continuity is (4.6), while one derives the following normal form for nonstandard integrability as in Example 2.1:

$$(\forall^{\text{st}} k')(\exists^{\text{st}} N')\big[(\forall \pi, \pi' \in P([0,1]))(\|\pi\|, \|\pi'\| \leq \tfrac{1}{N'} \to |S_\pi(f) - S_\pi(f)| \leq \tfrac{1}{k'})\big],$$

where $B(k', N', f)$ is the (internal) formula in square brackets and the underlined formula in (4.6) is abbreviated $A(f, N, k)$. Then $\mathsf{RIE}_{\mathsf{ns}}$ is the following implication

$$(\forall f : \mathbb{R} \to \mathbb{R})\underline{\big[(\forall^{\text{st}} k)(\exists^{\text{st}} N)A(f, N, k) \to (\forall^{\text{st}} k')(\exists^{\text{st}} N')B(k', N', f)\big]}. \tag{4.21}$$

and applying Remark 4.8, i.e. the step from (4.14) to (4.17), to the underlined formula in (4.21), we obtain the following:

$$(\forall f : \mathbb{R} \to \mathbb{R})(\forall^{\text{st}} g, k')(\exists^{\text{st}} N')\big[(\forall k)A(f, g(k), k) \to B(k', N', f)\big]. \tag{4.22}$$

Now, (4.22) is not a normal form, but note that it trivially implies

$$(\forall^{\text{st}} g, k')\underline{(\forall f : \mathbb{R} \to \mathbb{R})(\exists^{\text{st}} N')\big[(\forall k)A(f, g(k), k) \to B(k', N', f)\big]}, \tag{4.23}$$

where the underline formula has the right from to apply *Idealisation* I. We obtain:

$$(\forall^{\text{st}} g, k')(\exists^{\text{st}} w^{0^*})(\forall f : \mathbb{R} \to \mathbb{R})(\exists N' \in w)\big[(\forall k)A(f, g(k), k) \to B(k', N', f)\big],$$

which is a normal form. Define $N_0 := \max_{i<|w|} w(i)$ and note that

$$(\forall^{\text{st}} g, k')(\exists^{\text{st}} N_0)(\forall f : \mathbb{R} \to \mathbb{R})\big[(\forall k)A(f, g(k), k) \to B(k', N_0, f)\big], \tag{4.24}$$

since $B(k', M, f) \to B(k', K, f)$ for $K > M$ by the definition of Riemann integration. Applying Theorem 4.2 to '$\mathsf{P} \vdash (4.24)$', we obtain a term s such that

$$(\forall g, k')(\exists N_0 \in s(g, k'))(\forall f : \mathbb{R} \to \mathbb{R})\big[(\forall k)A(f, g(k), k) \to B(k', N_0, f)\big],$$

and defining $t(g, k')$ as $\max_{i<|s(g,k')|} s(g, k')(i)$, we obtain $\mathsf{RIE}_{\mathsf{ef}}(t)$ in light of the definitions of A, B from the beginning of this proof. $\square$

Note that the actual computation in $\mathsf{RIE}_{\mathsf{ef}}(t)$ only takes place on the modulus g; this is a consequence of $\mathsf{RIE}_{\mathsf{ns}}$ applying to *all* functions, not just the standard ones. Indeed, in the previous proof we could obtain (4.23) and then 'pull the standard existential quantifier $(\exists^{\text{st}} N)$ through the internal one $(\forall f)$' thanks to *Idealisation*, and then obtain (4.24) in which the quantifier $(\forall f)$ has no influence anymore on term extraction as in Theorem 4.2.

Clearly, the same strategy as in the previous proof works for *any* normal form in the scope of an internal quantifier. Indeed $(\forall z)(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y^\tau)\varphi(x,y,z)$ implies $(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} w^{\tau^*})(\forall z)(\exists y \in w)\varphi(x,y,z)$ by *Idealisation*. Furthermore, as is done in the previous proof for $\tau = 0$, if φ is (somehow) monotone in y, we may define y_0 as (some kind of) maximum of $w(i)$ for $i < |w|$, and obtain $(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y_0^\tau)(\forall z)\varphi(x,y,z)$. We finish this section with an even stronger result: a normal form in the scope of an *external* quantifier over all infinitesimals is again a normal form. This result also sheds light on the correctness of the intuitive infinitesimal calculus.

Theorem 4.12 (P). *For internal φ, $(\forall \varepsilon \approx 0)(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y^\tau)\varphi(x,y,\varepsilon)$ is equivalent to a normal form. If $\tau = 0$ and $(\forall z^0, w^0)\big(z > w \to (\varphi(x,w,\varepsilon) \to \varphi(x,z,\varepsilon))\big)$, then*

$$
\begin{aligned}
(\forall \varepsilon \approx 0)(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y^0)\varphi(x,y,\varepsilon) &\leftrightarrow (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y^0)(\forall \varepsilon \approx 0)\varphi(x,y,\varepsilon) \\
&\leftrightarrow (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y^0, N^0)(\forall \varepsilon)\big[|\varepsilon| < \tfrac{1}{N} \to \varphi(x,y,\varepsilon)\big].
\end{aligned}
\tag{4.25}
$$

Proof. Written out in full, the initial formula from the theorem is:

$$(\forall \varepsilon)\big[(\forall^{\mathrm{st}} k^0)(|\varepsilon| < \tfrac{1}{k}) \to (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y^\tau)\varphi(x,y,\varepsilon)\big],$$

and bringing outside all standard quantifiers as far as possible:

$$(\forall^{\mathrm{st}} x)\underline{(\forall \varepsilon)(\exists^{\mathrm{st}} y^\tau, k^0)\big[|\varepsilon| < \tfrac{1}{k} \to \varphi(x,y,\varepsilon)\big]},$$

the underlined formula is suitable for *Idealisation*. Applying the latter yields

$$(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} w^{0^*}, z^{\tau^*})(\forall \varepsilon)(\exists y^\tau \in z, k^0 \in w)\big[|\varepsilon| < \tfrac{1}{k} \to \varphi(x,y,\varepsilon)\big],$$

and let N^0 be the maximum of all $w(i)$ for $i < |w|$. We obtain:

$$(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} z^{\tau^*}, N)(\forall \varepsilon)(\exists y^\tau \in z)\big[|\varepsilon| < \tfrac{1}{N} \to \varphi(x,y,\varepsilon)\big].$$

If $\tau = 0$ and φ is monotone as above, let y_0 be $\max_{i<|z|} z(i)$ and obtain:

$$(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y_0, N)(\forall \varepsilon)\big[|\varepsilon| < \tfrac{1}{N} \to \varphi(x,y_0,\varepsilon)\big], \tag{4.26}$$

which is as required by (4.25) when pushing $(\exists^{\mathrm{st}} N)$ inside the square brackets. □

The monotonicity condition on φ in the theorem occurs a lot in analysis: In any 'epsilon-delta' definition, the formula resulting from removing the 'epsilon' and 'delta' quantifiers is monotone (in the sense of the theorem) in the 'delta' variable.

Furthermore, (4.25) explains why correct results can be produced by the intuitive infinitesimal calculus (from physics and engineering): Many arguments in the latter calculus produce formulas in the same form as the left-hand side of (4.25). Applying term extraction on the bottom normal form of (4.25), we obtain a term t such that:

$$(\forall x)(\exists y^0, N^0 \le t(x))(\forall \varepsilon)\big[|\varepsilon| < \tfrac{1}{N} \to \varphi(x,y,\varepsilon)\big], \tag{4.27}$$

which does not involve Nonstandard Analysis. However, assuming x as in (4.27) is only used in a certain (discreet) range, as seems typical of physics and engineering, we can fix *one* 'very large compared to the range of x' value N_0 for N and still obtain correct results in that $(\exists y)\varphi(x,y,1/N_0)$ for x in the prescribed range.

In short: (4.25) and (4.27) explain why the practice 'replacing infinitesimals by very small numbers' typical of physics and engineering can produce correct results.

4.3. **An example involving *Transfer*.**

4.3.1. *Introduction.* In the previous section, we provided examples of local constructivity, i.e. we stripped the nonstandard proofs of *Transfer* and *Standardisation* and obtained effective results using Theorem 4.2. A natural question, discussed in this section for *Transfer*, is what happens if we *do not omit* these axioms. As we will see, nonstandard proofs involving *Transfer* give rise to *relative computability results*.

By way of an elementary example, we study a nonstandard proof *involving Transfer* of the *intermediate value theorem* in Section 4.3.2. By way of an advanced example, we obtain in Section 4.3.3 effective results à la *Reverse Mathematics*[17]. In particular, we establish that P proves the equivalence between a fragment of *Transfer* and a nonstandard version of the *monotone convergence theorem* (involving nonstandard convergence), in line with the equivalences from Reverse Mathematics. From this nonstandard equivalence, we extract terms of Gödel's T which -intuitively speaking- convert a solution to the Halting problem into a solution to the monotone convergence theorem, and vice versa.

The above results, published first in part in [81], serve a dual purpose as follows:

(1) *Transfer* is shown to be *fundamentally non-constructive*, as it converts to Feferman's non-constructive mu-operator, i.e. essentially the Turing jump.
(2) Nonstandard equivalences as in Theorem 4.16 give rise to *effective* equivalences like (4.36) which are *rich in computational content*.

The first item directly vindicates local constructivity from Section 3.4. The second item is relevant as there is a field called *constructive Reverse Mathematics* (See e.g. [47] for an overview) where equivalences are proved in a system *based on intuitionistic logic*. In particular, the second item establishes that Reverse Mathematics done in classical Nonstandard Analysis gives rise to 'rather[18] constructive' Reverse Mathematics, in line with the observations from Section 4.1 regarding the semi-constructive status of classical Nonstandard Analysis.

We again stress that our results are only meant to illustrate the limits and potential of local constructivity. While e.g. the effective equivalences in Section 4.3.3 are not necessarily new or surprising, *it is surprising* that we can obtain them from nonstandard proofs *without any attempt whatsoever at working constructively.*

4.3.2. *The intermediate value theorem.* We study a nonstandard proof involving *Transfer* of the *intermediate value theorem* from Section 4.2.1. To this end, let $\mathsf{IVT}'_{\mathsf{ns}}$ be $\mathsf{IVT}_{\mathsf{ns}}$ with '$\approx$' replaced by '$=_{\mathbb{R}}$' in the conclusion. For the latter replacement, the following restriction of Nelson's axiom *Transfer* is needed:

$$(\forall^{\text{st}} f^1)\big[(\forall^{\text{st}} n^0) f(n) \neq 0 \to (\forall m) f(m) \neq 0\big]. \qquad (\Pi^0_1\text{-}\mathsf{TRANS})$$

We also need Feferman's *non-constructive mu-operator* ([5]) as follows:

$$(\exists \mu^2)\big[(\forall f^1)((\exists n) f(n) = 0 \to f(\mu(f)) = 0)\big], \qquad (\mu^2)$$

where $\mathsf{MU}(\mu)$ is the formula in square brackets in (μ^2). Finally, let $\mathsf{IVT}_{\mathsf{ef}}(t)$ state that $t(f,g)$ is the intermediate value for f with g its modulus of continuity.

---

[17]We refer to [86, 87] for an overview of Friedman's foundational program *Reverse Mathematics*, first introduced in [33, 34].

[18]Again, we do not claim that (4.36) counts as constructive mathematics (due to the classical base theory $\mathsf{E}\text{-}\mathsf{PA}^\omega$), but the former formula contains too much (automatically extracted from the nonstandard equivalence) computational information to be dismissed as 'non-constructive'.

Theorem 4.13. *From the proof '$\mathsf{P} \vdash \Pi_1^0\text{-}\mathsf{TRANS} \to \mathsf{IVT}_{\mathsf{ns}}$', we obtain a term t such that $\mathsf{E\text{-}PA}^\omega$ proves $(\forall \mu^2)(\mathsf{MU}(\mu) \to \mathsf{IVT}_{\mathsf{ef}}(t))$.*

Proof. Clearly, Π_1^0-TRANS is equivalent to the following normal form:

$$(\forall^{\text{st}} g^1)(\exists^{\text{st}} m)\big[(\exists n^0)g(n) = 0 \to (\exists i \leq m)g(i) = 0\big], \tag{4.28}$$

while $\mathsf{IVT}'_{\mathsf{ns}}$ has a normal form $(\forall^{\text{st}} f, g)(\exists^{\text{st}} x)C(f, g, x)$, where $C(f, g, x)$ expresses that if g is a modulus of uniform continuity of $f \in D$, then $f(x) =_{\mathbb{R}} 0$. The latter normal form is obtained in the same way as for $\mathsf{IVT}_{\mathsf{ns}}$ in the proof of Theorem 4.7. Now let $D(g, m)$ be the formula in square brackets in (4.28) and note that Π_1^0-$\mathsf{TRANS} \to \mathsf{IVT}_{\mathsf{ns}}$ implies the following implication between normal forms:

$$(\forall^{\text{st}} g)(\exists^{\text{st}} m)D(g, m) \to (\forall^{\text{st}} f, g)(\exists^{\text{st}} x)C(f, g, x).$$

Following Remark 4.8, we may conclude that P proves:

$$(\forall^{\text{st}} f, g, \mu^2)(\exists^{\text{st}} x)\big[(\forall g)D(g, \mu(g)) \to C(f, g, x)\big],$$

and applying Theorem 4.2, we obtain a term s such that $\mathsf{E\text{-}PA}^\omega$ proves:

$$(\forall f, g, \mu^2)(\exists x \in s(f, g, \mu))\big[(\forall g)D(g, \mu(g)) \to C(f, g, x)\big]. \tag{4.29}$$

We recognise the antecedent of (4.29) as (a slight variation of) $\mathsf{MU}(\mu)$ and note that the implication in (μ^2) is also an equivalence. In other words, Feferman's mu-operator allows us to decide Π_1^0-formulas, which includes formulas like $f(x) =_{\mathbb{R}} 0$. Thus, given $(\exists x \in s(f, g, \mu))(f(x) =_{\mathbb{R}} 0)$ as in (4.29), we just use the mu-operator to test which $s(f, g, \mu)(i)$ for $i < |s(f, g, \mu)|$ is an intermediate value of f. Let the term t be such that $t(f, g, \mu)$ is such an intermediate value and note that (4.29) implies $(\forall \mu^2)(\mathsf{MU}(\mu) \to \mathsf{IVT}_{\mathsf{ef}}(t))$, i.e. we are done. □

In conclusion, nonstandard proofs involving *Transfer* give rise to *relative computability results* such as in the previous theorem. The result in the latter is not meant to be new or optimal (See [55] for such results), but merely illustrate *why* we remove *Transfer* as dictated by local constructivity: Even a small fragment like Π_1^0-TRANS gives rise to (what is essentially) a solution to the Halting problem.

The conclusion of Theorem 4.13 need also not be surprising: IVT is rejected in constructive mathematics and it hence stands to reason that we need some non-computable object, like e.g. Feferman's mu-operator, to compute intermediate values. In fact, Kohlenbach proves versions of $(\mu^2) \leftrightarrow (\exists \Phi)\mathsf{IVT}_{\mathsf{ef}}(\Phi)$ in [55, §3].

4.3.3. *The monotone convergence theorem.* In this section, we extract (effective) equivalences à la Reverse Mathematics from (non-effective) equivalences in Nonstandard Analysis involving *Transfer*.

Firstly, we introduce a nonstandard version of the monotone convergence theorem MCT involving *nonstandard convergence*, as follows:

Definition 4.14. [$\mathsf{MCT}_{\mathsf{ns}}$] For every *standard* sequence x_n of reals, we have

$$(\forall n \in \mathbb{N})(0 \leq_{\mathbb{R}} x_n \leq_{\mathbb{R}} x_{n+1} \leq_{\mathbb{R}} 1) \to (\forall N, M \in \mathbb{N})[\neg\text{st}(N) \wedge \neg\text{st}(M) \to x_M \approx x_N].$$

Note that the conclusion expresses the nonstandard convergence of x_n. The corresponding *effective/constructive*[19] version is:

[19]Similar to the case of continuity, a modulus of convergence is part and parcel of the constructive definition of convergence, as discussed in e.g. [11, p. 27].

Definition 4.15. [$\mathsf{MCT}_{\mathsf{ef}}(t)$] For any sequence x_n of reals, we have
$(\forall n \in \mathbb{N})(0 \leq_{\mathbb{R}} x_n \leq_{\mathbb{R}} x_{n+1} \leq_{\mathbb{R}} 1) \rightarrow (\forall k, N, M \in \mathbb{N})[N, M \geq t(x_{(\cdot)})(k)) \rightarrow |c_M - c_N| \leq \frac{1}{k}]$.

We also require the following functional, which is equivalent to (μ^2) by ([55, §3]),

$$(\exists \varphi^2)\big[(\forall f^1)((\exists n)f(n) = 0 \leftrightarrow \varphi(f) = 0)\big], \tag{$\exists^2$}$$

and which is called the *Turing jump functional*. Denote by $\mathsf{TJ}(\varphi)$ the formula in square brackets in $(\exists^2)$. We have the following theorem and corollary.

Theorem 4.16. *The system* P *proves that* $\mathsf{MCT}_{\mathsf{ns}} \leftrightarrow \Pi^0_1\text{-}\mathsf{TRANS}$.

Proof. First of all, to establish $\mathsf{MCT}_{\mathsf{ns}} \rightarrow \Pi^0_1\text{-}\mathsf{TRANS}$, fix standard f^1 such that $(\forall^{\mathrm{st}} n)f(n) \neq 0$ and define the *standard* sequence $x_{(\cdot)}$ of reals as follows

$$x_k := \begin{cases} 0 & (\forall i \leq k)f(i) \neq 0 \\ \sum_{i=1}^{k} \frac{1}{2^i} & \text{otherwise} \end{cases}. \tag{4.30}$$

Clearly, $x_{(\cdot)}$ is weakly increasing and hence nonstandard convergent by $\mathsf{MCT}_{\mathsf{ns}}$. However, if $(\exists m_0)f(m_0+1) = 0$ and m_0 is the least such number, we would have $0 = x_{m_0} \not\approx x_{m_0+1} \approx 1$. The latter contradicts $\mathsf{MCT}_{\mathsf{ns}}$ and we must therefore have $(\forall n^0)(f(n) \neq 0)$. Thus, $\Pi^0_1\text{-}\mathsf{TRANS}$ follows and we now prove the other direction in two different ways. To this end, let x_n be as in $\mathsf{MCT}_{\mathsf{ns}}$ and consider the formula

$$(\forall^{\mathrm{st}} k \in \mathbb{N})(\exists^{\mathrm{st}} m \in \mathbb{N})\underline{(\forall^{\mathrm{st}} N, M \geq m)[|x_M - x_N| \leq \tfrac{1}{k}]}, \tag{4.31}$$

which expresses that x_n 'epsilon-delta' converges relative to 'st'. If (4.31) is false, there is some k_0 such that $(\forall^{\mathrm{st}} m \in \mathbb{N})(\exists^{\mathrm{st}} M > N \geq m)[x_M > x_N + \frac{1}{k_0}]$. However, applying[20] the latter formula k_0+1 times, we obtain some M_0 such that $x_{M_0} >_{\mathbb{R}} 1$, contradiction our assumptions. Hence, (4.31) holds and applying $\Pi^0_1\text{-}\mathsf{TRANS}$ to the innermost underlined formula yields that x_n is nonstandard convergent.

For a second proof of $\Pi^0_1\text{-}\mathsf{TRANS} \rightarrow \mathsf{MCT}_{\mathsf{ns}}$ based on the existing literature, note that $\Pi^0_1\text{-}\mathsf{TRANS}$ implies the following normal form:

$$(\forall^{\mathrm{st}} f^1)(\exists^{\mathrm{st}} m)\big[(\exists n^0)f(n) = 0 \rightarrow (\exists i \leq m)f(i) = 0\big], \tag{4.32}$$

to which $\mathsf{HAC}_{\mathsf{int}}$ may be applied to obtain standard $\nu^{1\rightarrow 0^*}$ such that

$$(\forall^{\mathrm{st}} f^1)(\exists m \in \nu(f))\big[(\exists n^0)f(n) = 0 \rightarrow (\exists i \leq m)f(i) = 0\big]. \tag{4.33}$$

Now define $\mu(f)$ as the maximum of $\nu(f)(i)$ for $i < |\nu(f)|$ and conclude that

$$(\forall^{\mathrm{st}} f^1)\big[(\exists n^0)f(n) = 0 \rightarrow (\exists i \leq \mu(f))f(i) = 0\big]. \tag{4.34}$$

Note that we actually have equivalence in (4.34), i.e. μ allows us to decide existential formulas as long as a standard function describes the quantifier-free part as in the antecedent of (4.34). Hence, the system $\mathsf{P} + \Pi^0_1\text{-}\mathsf{TRANS}$ proves arithmetical comprehension ACA_0 (See [87, III]) relative to the 'st'. A well-known result from Reverse Mathematics (See [87, III.2.2]) is that RCA_0 proves $\mathsf{ACA}_0 \leftrightarrow \mathsf{MCT}$. Since P also proves the axioms of RCA_0 relative to 'st', the system $\mathsf{P} + \Pi^0_1\text{-}\mathsf{TRANS}$ proves $\mathsf{MCT}^{\mathrm{st}}$. Hence, for x_n as in the latter, there is standard $x \in \mathbb{R}$ such that

$$(\forall^{\mathrm{st}} k \in \mathbb{N})(\exists^{\mathrm{st}} m \in \mathbb{N})\underline{(\forall^{\mathrm{st}} N, M \geq m)[|x_M - x| \leq \tfrac{1}{k}]}. \tag{4.35}$$

[20]To 'apply this formula k_0+1 times', replace all x_n in $(\forall^{\mathrm{st}} m \in \mathbb{N})(\exists^{\mathrm{st}} M > N \geq m)[x_M > x_N + \frac{1}{k_0}]$ by approximations up to some fixed nonstandard number and apply $\mathsf{HAC}_{\mathsf{int}}$. Since all numbers are rationals (due to the approximation), we can test which number is the correct one.

Applying Π_1^0-TRANS to the underline formula in (4.35) now finishes the proof. $\square$

Corollary 4.17. *From any proof of* $\mathsf{MCT}_{\mathsf{ns}} \leftrightarrow \Pi_1^0$-$\mathsf{TRANS}$ *in* P*, two terms* s, u *can be extracted such that* $\mathsf{E\text{-}PA}^\omega$ *proves:*

$$(\forall \mu^2)\big[\mathsf{MU}(\mu) \to \mathsf{MCT}_{\mathsf{ef}}(s(\mu))\big] \wedge (\forall t^{1\to 1})\big[\mathsf{MCT}_{\mathsf{ef}}(t) \to \mathsf{MU}(u(t))\big]. \tag{4.36}$$

Proof. We prove the corollary for the implication Π_1^0-$\mathsf{TRANS} \to \mathsf{MCT}_{\mathsf{ns}}$ and leave the other one to the reader. The sentence $\mathsf{MCT}_{\mathsf{ns}}$ is readily converted to

$$(\forall^{\mathrm{st}} x_{(\cdot)}^{0\to 1}, k)(\exists^{\mathrm{st}} m)\big[(\forall n^0)(0 \le x_n \le x_{n+1} \le 1) \tag{4.37}$$
$$\to (\forall N, M \ge m)[|x_M - x_N| \le \tfrac{1}{k}],$$

in the same way as in Example 2.1. Now let A (resp. B) be the formula in square brackets in (4.34) (resp. (4.37)) and note that the existence of standard μ^2 as in (4.34) implies Π_1^0-TRANS. Hence, Π_1^0-$\mathsf{TRANS} \to \mathsf{MCT}_{\mathsf{ns}}$ gives rise to the following:

$$\big[(\exists^{\mathrm{st}} \mu^2)(\forall^{\mathrm{st}} f^1)A(f, \mu(f))\big] \to (\forall^{\mathrm{st}} x_{(\cdot)}^{0\to 1}, k^0)(\exists^{\mathrm{st}} m^0)B(x_{(\cdot)}, k, m), \tag{4.38}$$

and bringing outside the standard quantifiers, we obtain

$$(\forall^{\mathrm{st}} x_{(\cdot)}^{0\to 1}, k^0, \mu^2)(\exists^{\mathrm{st}} m, f)[A(f, \mu(f)) \to B(x_{(\cdot)}, k, m)], \tag{4.39}$$

which is a normal form as the formula in square brackets is internal. Now applying Theorem 4.2 to '$\mathsf{P} \vdash$ (4.39)' yields a term t such that $\mathsf{E\text{-}PA}^\omega$ proves

$$(\forall x_{(\cdot)}^{0\to 1}, k^0, \mu^2)(\exists m, f \in t(\mu, x_{(\cdot)}, k))[A(f, \mu(f)) \to B(x_{(\cdot)}, k, m)], \tag{4.40}$$

and define $s(\mu, x_{(\cdot)}, k)$ to be the maximum of all entries for m in $t(\mu, x_{(\cdot)}, k)$. We immediately obtain, using classical logic, that

$$(\forall \mu^2)[(\forall f^1)A(f, \mu(f)) \to (\forall x_{(\cdot)}^{0\to 1}, k^0)B(x_{(\cdot)}, k, s(\mu, x_{(\cdot)}, k))],$$

which is exactly as required by the theorem, i.e. the first conjunct of (4.36). $\square$

With some effort, the proof of $\mathsf{MCT}_{\mathsf{ns}} \leftrightarrow \Pi_1^0$-$\mathsf{TRANS}$ goes through in H, i.e. the theorem is constructive, and the corollary then applies to H and $\mathsf{E\text{-}HA}^\omega$. The axioms from Definition A.7 are crucial in this context. In any case, the previous results indicate that *Transfer* is *highly non-constructive*, as it is converted into the Turing jump (or stronger) after term extraction, thus lending credence to the **local constructivity** of Nonstandard Analysis.

Furthermore, we stress that the results in (4.36) are not (necessarily) surprising in and of themselves. *What is surprising* is that we can 'algorithmically' derive these effective results from the *quite simple* classical proof of Π_1^0-$\mathsf{TRANS} \leftrightarrow \mathsf{MCT}_{\mathsf{ns}}$, in which *no efforts towards effective results are made.* Furthermore, while (4.36) deals with higher types, it is possible to derive *classical* computability theory, i.e. dealing only with subsets of $\mathbb{N}$, as discussed in Section 5.4.

Finally, we point out one entertaining property of *Transfer* with regard to the law of excluded middle: When working in P, there are in general *three* possibilities for any internal formula $A(x)$: (i) $A(x)$ holds for all x, (ii) there is *standard* x such that $\neg A(x)$, and (iii) there is y such that $\neg A(y)$, but no such standard number exists. Hence, *in the extended language of* P there is a **third**[21] possibility alongside of the two usual ones, namely that there is a counterexample, but it is not standard.

[21] Note that one can prove using intuitionistic logic that $\neg(\neg Q \wedge \neg\neg Q)$ for any Q, i.e. there is no explicit 'third option' in which both Q and $\neg Q$ are false ([18, 36]).

Now, the *Transfer* axiom T excludes the aforementioned third possibility, i.e. in IST we have $(\exists^{\text{st}}x)\neg A(x) \vee (\forall x)A(x)$. Similar to Section 4.1, we observe that the fundamental change from mainstream Peano arithmetic $\mathsf{E\text{-}PA}^{\omega}$ to the system of Nonstandard Analysis P has an influence on the associated logic: The 'usual' law of excluded middle included in P allows for three possibilities as above, while *Transfer* reduces this spectrum to two, i.e. the latter plays the role of the law of excluded middle *in the extended language of Nonstandard Analysis*.

4.4. **An example involving *Standard Part*.** In Section 4.2, we provided examples of local constructivity, i.e. we stripped the nonstandard proofs of *Transfer* and *Standardisation* and obtained effective results using Theorem 4.2. A natural question is what happens if we *do not omit* the latter axiom.

To this end, we study in Section 4.4.1 a nonstandard proof *involving Standard Part* of RIE formulated with *pointwise* continuity. As we will see, nonstandard proofs involving *Standard Part* give rise to *rather exotic* relative computability results. In particular, similar to Section 4.3.3, we establish that P proves the equivalence between STP, a fragment of *Standard Part*, and a nonstandard version of RIE. Applying Theorem 4.2 to this nonstandard equivalence, we observe that STP is converted to the *special fan functional*, a cousin of Tait's fan functional with rather exotic properties, as discussed in Sections 4.4.1 and 4.4.2.

The above results serve a triple purpose as follows:

(1) *Standard Part* is *non-constructive*, as it gives rise to the *special fan functional*, which is extremely hard to compute in classical mathematics.
(2) The *combination* of the fragments of *Standard Part* and *Transfer* corresponding to WKL_0 and ACA_0 gives rise to $\Pi^1_1\text{-}\mathsf{CA}_0$. Thus, *Standard Part* is 'even more' non-constructive in the presence of *Transfer*.
(3) Nonstandard equivalences as in Theorem 4.19 give rise to *effective* equivalences like in Theorem 4.21 which are *rich in computational content*.

The first and second item directly vindicate local constructivity from Section 3.4. The final item is relevant for *constructive Reverse Mathematics* as discussed in the previous section. By contrast, we discuss in Section 4.4.3 how results involving the special fan functional give rise to 'normal' results involving known objects.

4.4.1. *Weak König's lemma and continuity.* In this section, we study a nonstandard proof involving a fragment of *Standard Part*. In particular, we study the theorem $\mathsf{RIE}^{\mathsf{pw}}$ that *a pointwise continuous function on the unit interval is integrable*.

First of all, *weak König's lemma* (WKL) is the statement *every infinite binary tree has a path* and gives rise to the second 'Big Five' system of Reverse Mathematics (See e.g. [87, IV]). The following fragments of *Standard Part* are equivalent (See Theorem A.14) and constitute the nonstandard counterpart of weak König's lemma.

Definition 4.18. [Equivalent fragments of *Standard Part*]

$$(\forall \alpha^1 \leq_1 1)(\exists^{\text{st}}\beta^1 \leq_1 1)(\alpha \approx_1 \beta). \tag{STP}$$

$$(\forall x \in [0,1])(\exists^{\text{st}} y \in [0,1])(x \approx y). \tag{$\mathsf{STP}_{\mathbb{R}}$}$$

$$(\forall f^1)(\exists^{\text{st}} g^1)\big((\forall^{\text{st}} n^0)(\exists^{\text{st}} m^0)(f(n) = m) \rightarrow f \approx_1 g\big). \tag{4.41}$$

$$(\forall T^1 \leq_1 1)\big[(\forall^{\text{st}} n)(\exists \beta^0)(|\beta| = n \wedge \beta \in T) \rightarrow (\exists^{\text{st}}\alpha^1 \leq_1 1)(\forall^{\text{st}} n^0)(\overline{\alpha} n \in T)\big], \tag{4.42}$$

$$(\forall^{\rm st} g^2)(\exists^{\rm st} w)\big[(\forall T^1\leq_1 1)\big((\forall \alpha^1\in w(1))(\overline{\alpha}g(\alpha)\not\in T) \rightarrow (\forall \beta\leq_1 1)(\exists i\leq w(2))(\overline{\beta}i\not\in T)\big)\big]. \tag{4.43}$$

There is no deep philosophical meaning in the words 'nonstandard counterpart': This is just what the principle STP is called in the literature (See e.g. [88]). Note that (4.42) is just $\mathsf{WKL}^{\rm st}$ with the 'st' dropped in the leading quantifier, while (4.41) is a version of the axiom of choice limited to Π_1^0-formulas (See [87, Table 4, p. 54]), and (4.43) is a normal form for STP. Also, $\mathsf{STP}_{\mathbb{R}}$ expresses the nonstandard compactness of the unit interval by [45, p. 42]. Thus, these nonstandard equivalences reflect the 'usual' equivalences involving WKL from [87, IV].

Now, by [87, IV.2.7], WKL is equivalent to $\mathsf{RIE}^{\mathsf{pw}}$ and the nonstandard counterparts behave in exactly the same way: letting $\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ns}}$ be the statement *every* (*pointwise*) *nonstandard continuous function is nonstandard integrable on the unit interval*, we have the following (expected) theorem.

Theorem 4.19. *The system* P *proves that* $\mathsf{STP}_{\mathbb{R}}\leftrightarrow \mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ns}}$.

Proof. For the forward implication, in light of the definitions of (uniform) nonstandard continuity (4.1) and (4.5), the latter follows from the former by $\mathsf{STP}_{\mathbb{R}}$, and $\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ns}}$ follows from Theorem 4.10 and noting that the latter establishes that $\mathsf{RIE}_{\mathsf{ns}}$ is provable in P. For the reverse implication, assume $\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ns}}$ and suppose $\mathsf{STP}_{\mathbb{R}}$ is false, i.e. there is $x_0\in[0,1]$ such that $(\forall^{\rm st} y\in[0,1])(x\not\approx y)$. Note that $0\ll x_0\ll 1$ as the endpoints of the unit interval are standard. Fix nonstandard N^0 and define the function $\sin\big(\frac{1}{|x-x_0|+\frac{1}{N}}\big)$. The latter is well-defined for any $x\in[0,1]$ and nonstandard pointwise continuous, but not nonstandard continuous at x_0. □

While Theorem 4.19 may be expected in light of known Reverse Mathematics results, applying term extraction as in Theorem 4.2 to the former theorem, we obtain an object with *unexpected* properties, namely the *special fan functional*.

Definition 4.20. [Special fan functional] We define $\mathsf{SCF}(\Theta)$ as follows for $\Theta^{(2\rightarrow(0\times 1^*))}$:

$$(\forall g^2, T^1\leq_1 1)\big[(\forall \alpha\in\Theta(g)(2))(\overline{\alpha}g(\alpha)\not\in T)\rightarrow(\forall\beta\leq_1 1)(\exists i\leq \Theta(g)(1))(\overline{\beta}i\not\in T)\big].$$

Any functional Θ satisfying $\mathsf{SCF}(\Theta)$ is referred to as a *special fan functional*.

Note that there is *no unique* special fan functional, i.e. it is in principle incorrect to make statements about 'the' special fan functional. Let $\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ef}}(t)$ be the statement that $t(g)$ is a modulus of Riemann integration for f if g is a modulus of pointwise continuity for f. We have the following theorem.

Theorem 4.21. *From any proof* $\mathsf{P}\vdash[\mathsf{STP}\rightarrow\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ns}}]$*, a term t can be extracted such that* $\mathsf{E\text{-}PA}^{\omega}$ *proves* $(\forall\Theta)[\mathsf{SCF}(\Theta)\rightarrow\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ef}}(t(\Theta))]$.

Proof. A normal form for STP is given by (4.43) while a normal form for $\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ns}}$ is similar to the normal for (4.24) for $\mathsf{RIE}_{\mathsf{ns}}$. The rest of the proof is now straightforward in light of Remark 4.8. □

The result of the previous theorem is similar to what one would expect from the previous sections. However, the special fan functional has rather *exotic computational properties*, as we discuss in the next section.

Finally, we point out one entertaining property of *Standard Part* with regard to the law of excluded middle: When working in P, there are in general *three* possibilities for any internal formula $A(x)$: (i) $A(x)$ holds for all x, (ii) there is *standard* x such that $\neg A(x)$, and (iii) there is y such that $\neg A(y)$, but no such standard number exists. Hence, *in the extended language of* P there is a **third**[22] possibility alongside of the two usual ones, namely that there is a counterexample, but it is not standard. Now, the *Standard Part* axiom S excludes the latter third possibility in certain cases, i.e. in $\mathsf{P}+\mathsf{STP}$ we have $(\exists^{\mathrm{st}}x^1 \leq_1 1)\neg A(x) \vee (\forall x^1 \leq_1 1)A(x)$ assuming A 'is nonstandard continuous' as follows:

$$(\forall^{\mathrm{st}} f^1)(\forall g^1)(f \approx_1 g \rightarrow (A(g) \rightarrow A(f))), \tag{4.44}$$

which expresses[23] that A only depends on some (standard) initial segment of a (standard) witness, similar to Brouwer's axiom $\mathsf{WC\text{-}N}$ ([4]). Formulas behaving as in (4.44) are readily found in computability theory; take for example $A(f) \equiv (\varphi^f_{e,s}(n) = m)$ with standard numerical parameters. Thus, the following typical example of *Transfer* also follows from STP thanks to the above considerations:

$$(\forall^{\mathrm{st}} e,s,m,n)\big[(\exists B^1)(\varphi^B_{e,s}(n)=m) \rightarrow (\exists^{\mathrm{st}} B^1)(\varphi^B_{e,s}(n)=m)\big].$$

In conclusion, as in Section 4.1 and the previous one, we observe that the fundamental change from mainstream Peano arithmetic $\mathsf{E\text{-}PA}^\omega$ to the system of Nonstandard Analysis P has an influence on the associated logic: The 'usual' law of excluded middle included in P allows for three possibilities as above, while *Standard Part* reduces this spectrum to two in certain cases, i.e. the latter can play the role of the law of excluded middle *in the extended language of Nonstandard Analysis.*

4.4.2. *Computational properties of the special fan functional.* We discuss the exotic computational properties of the special fan functional. As it turns out, the latter is *extremely hard to compute* in classical mathematics, while easy to compute in intuitionistic mathematics. All results on the special fan functional are from [65, 78].

First of all, the 'usual' fan functional Φ^3 was introduced by Tait (See e.g. [59, 66]) as an example of a functional not computable (in the sense of Kleene's S1-S9; [59, 66]) over the total continuous functionals. The usual fan functional Φ returns a modulus of uniform continuity $\Phi(Y)$ on Cantor space on input a continuous functional Y^2. Since Cantor space and the unit interval are isomorphic, it is clear that the usual fan functional Φ also computes a modulus of Riemann integration for (pointwise) continuous functionals on the unit interval. With some effort, these kind of results follow from the implication $\mathsf{WKL}^{\mathrm{st}} \rightarrow (\mathsf{RIE}^{\mathsf{pw}})^{\mathrm{st}}$. In particular, $\mathsf{WKL}^{\mathrm{st}}$ is translated to (something similar to) the usual fan functional after term extraction as in Theorem 4.2. Furthermore, the usual fan functional may be computed (via a term in Gödel's T) from (μ^2) by combining the results in [56, 57].

Secondly, the special fan functional fulfils a role similar to the usual fan functional in Theorem 4.21: it computes a modulus of Riemann integration for f from a modulus of pointwise continuity from f. However, (μ^2) cannot compute (in the sense of Kleene's S1-S9) any Θ as in $\mathsf{SCF}(\Theta)$; the same holds for *any* type two functional, including the *Suslin functional* which corresponds to $\Pi^1_1\text{-}\mathsf{CA}_0$, the strongest Big Five system ([55]). Thus, it seems one needs the functional $(\mathcal{E}_2)$:

$$(\exists \xi^3)(\forall Y^2)\big[(\exists f^1)(Y(f)=0) \leftrightarrow \xi(Y)=0\big], \tag{$\mathcal{E}_2$}$$

[22] Recall Footnote 21 concerning the truth values of intuitionistic logic.

[23] To see this, resolve '$\approx_1$' in (4.44) and push the resulting extra standard quantifier to the front using *idealisation*.

to compute Θ such that $\mathsf{SCF}(\Theta)$. But $(\mathcal{E}_2)$ implies *second-order arithmetic*, which is *off-the-chart* qua computational strength compared to (μ^2), i.e. the special fan functional is extremely *hard to compute* in classical mathematics. The fact that (μ^2) does not compute the special fan functional translates to the nonstandard fact that $\mathsf{P}+\Pi_1^0\text{-}\mathsf{TRANS}$ does not prove STP, in contrast to the situation for the standard counterparts ACA_0 and WKL_0 from Reverse Mathematics where the latter implies the former. Furthermore, the combination of STP and $\Pi_1^0\text{-}\mathsf{TRANS}$ allows us to derive a version of *Transfer* for Π_1^1-formulas. As a result, (μ^2) and the special fan functional together compute a version of the Suslin functional, i.e. $\Pi_1^1\text{-}\mathsf{CA}_0$.

Thirdly, by contrast, $\mathsf{E\text{-}PA}^\omega+(\exists\Theta)\mathsf{SCF}(\Theta)$ is a conservative extension of $\mathsf{E\text{-}PA}^\omega+\mathsf{WKL}$, i.e. the special fan functional has quite *weak first-order strength* in isolation. This conservation result requires the *intuitionistic fan functional* Ω^3 as follows:

$$(\forall Y^2)(\forall f^1,g^1\leq_1 1)(\overline{f}\Omega(Y)=_0\overline{g}\Omega(Y)\rightarrow Y(f)=_0 Y(g)) \qquad (\mathsf{MUC}(\Omega))$$

which results in a conservative extension of $\mathsf{E\text{-}PA}^\omega+\mathsf{WKL}$ by [55, Theorem 3.15]. Now, the intuitionistic fan functional computes the special one via a term of Gödel's T by Theorem 4.22 and its corollary. Thus, the special fan functional is rather *easy* to compute in intuitionistic mathematics. Proofs my be found in Section A.5.

Theorem 4.22. *The axiom* STP *can be proved in* P *plus the axiom*

$$(\forall^{\text{st}}Y^2)(\forall f^1,g^1\leq_1 1)(f\approx_1 g\rightarrow Y(f)=_0 Y(g)). \qquad (\mathsf{NUC})$$

In particular $\mathsf{P}+(\exists^{\text{st}}\Omega^3)\mathsf{MUC}(\Omega)$ *proves* STP.

Corollary 4.23. *From the proof in* P *that* $\mathsf{NUC}\rightarrow\mathsf{STP}$*, a term* $t^{3\rightarrow 3}$ *can be extracted such that* $\mathsf{E\text{-}PA}^\omega$ *proves* $(\forall\Omega^3)\big[\mathsf{MUC}(\Omega)\rightarrow\mathsf{SCF}(t(\Omega))\big]$.

We point out that Kohlenbach ([56, §5]) has introduced axiom schemas with properties similar to the special fan functional Θ. These schemas are logical in nature in that they are formulated using comprehension over formula classes. By contrast, Θ is 'purely mathematical' as it emerges from the normal form of STP.

Finally, we point out an anecdote by Friedman regarding Robinson from 1966.

> I remember sitting in Gerald Sacks' office at MIT and telling him about this [version of Nonstandard Analysis based on PA] and the conservative extension proof. He was interested, and spoke to A. Robinson about it, Sacks told me that A. Robinson was disappointed that it was a conservative extension. [35]

In light of the previous quote, we believe Robinson would have enjoyed learning about the 'new' mathematical object that is the special fan functional originating from Nonstandard Analysis. As it happens, many (if not most) theorems of second-order arithmetic can be modified to yield similar 'special' functionals with exotic computational properties.

In conclusion, we provided an example of a nonstandard proof involving *Standard Part* (STP in particular) which gives rise to a *rather exotic* relative computability result, namely involving the *special fan functional* which is extremely hard to compute classically, and rather easy to compute intuitionistically. This observation is particularly surprising as by (4.42) STP is so similar to WKL relative to 'st', and the latter behaves completely 'standard'.

4.4.3. *Non-exotic computability results and* Standard Part. As discussed at the end of Section 1, part of this paper's *Catch22* is that we cannot show too much 'strange stuff' emerging from Nonstandard Analysis lest the reader get the idea that Nonstandard Analysis is somehow fundamentally strange. For this reason, we discuss in this section how the above results involving the (admittedly strange) special fan functional give rise to 'normal' results with more computational content. For reasons of space, we only provide a sketch while technical details are in [65].

First of all, consider the following nonstandard version of the axiom of choice:

$$(\forall^{\mathrm{st}} n^0)(\exists^{\mathrm{st}} x^\rho)\Phi(n,x) \rightarrow (\exists^{\mathrm{st}} F^{0\rightarrow\rho})(\forall^{\mathrm{st}} n^0)\Phi(n,F(n)), \tag{$\mathsf{AC}^0_{\mathsf{ns}}$}$$

where Φ can be any formula in the language of P. As shown in [8, 31], the system $\mathsf{P}+\mathsf{AC}^0_{\mathsf{ns}}$ satisfies a version of Theorem 4.2 but with t defined using *bar recursion*, an advanced computation scheme which (somehow) embodies the computational content of the axiom of choice (See [54] for details on bar recursion). Hence, since $\mathsf{AC}^0_{\mathsf{ns}}$ readily implies $\mathsf{STP}_{\mathbb{R}}$ by Theorem A.14, we may consider the proof $\mathsf{P}+\mathsf{AC}^0_{\mathsf{ns}} \vdash \mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ns}}$. Instead of Theorem 4.21, we obtain a *bar recursive* term t such that $\mathsf{RIE}^{\mathsf{pw}}_{\mathsf{ef}}(t)$. While the special fan functional is gone, we now have a term depending on bar recursion, which seems somewhat overkill for the modest theorem $\mathsf{RIE}^{\mathsf{pw}}$.

Secondly, Kohlenbach discusses the ECF-translation for the base theory RCA^ω_0 of higher-order Reverse Mathematics from [55]. The original ECF-translation may be found in [97, §2.6.5, p. 141] and is somewhat technical in nature. Intuitively, the ECF-translation replaces all objects of type two or higher by *type one associates*, also known as *Reverse Mathematics codes*. In particular, as shown by Kohlenbach in [55, §3], if $\mathsf{RCA}^\omega_0 \vdash A$, then $\mathsf{RCA}^2_0 \vdash [A]_{\mathsf{ECF}}$, where A is any formula in the language of finite types, $[A]_{\mathsf{ECF}}$ is the ECF-translation of A, and RCA^2_0 is essentially the base theory RCA_0 of Reverse Mathematics. Serendipitously, the ECF-translation applied to the conclusion of Theorem 4.21 gives rise to (a statement equivalent to) $\mathsf{WKL} \rightarrow \mathsf{RIE}^{\mathsf{pw}}$ (See [65]). In other words, while STP translates to the (arguably strange) special fan functional, the latter becomes 'plain old' WKL when applying the ECF-translation! To use an imaginary proverb: All is standard that ends standard!

In conclusion, we have shown how results involving the (admittedly strange) special fan functional from the previous section give rise to 'normal' results. The ECF-approach is preferable over the one involving bar recursion, in our opinion.

4.5. **There and back again.** The examples of local constructivity discussed in Sections 4.2 to 4.4 are *by design* such that theorems in Nonstandard Analysis are converted to theorems of computable or constructive mathematics. As suggested by the title, it is then a natural question whether the latter also again imply the former, i.e. do theorems of computable mathematics also (somehow) imply theorems of Nonstandard Analysis? Are they (somehow) equivalent?

At the theoretical level, the answer is a resounding 'yes', as is clear from the following theorem which is the 'inverse' of term extraction as in Theorem 4.2.

Theorem 4.24. *Let φ be internal and t a term. The system P proves*

$$(\forall x)(\exists y \in t(x))\varphi(x,y) \rightarrow (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\varphi(x,y). \tag{4.45}$$

Furthermore, if $\mathsf{E}\text{-}\mathsf{PA}^\omega \vdash (\forall x)(\exists y \in t(x))\varphi(x,y)$, then $\mathsf{P} \vdash (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\varphi(x,y)$.

Proof. Both claims follow immediately by noting that every term t is standard in P by the basic axioms in Definition A.3, and that the latter also implies that $t(x)$ is standard for standard x. $\square$

At the practical level things are slightly more complicated: While normal forms follow from computational statements like in (4.45), these normals forms are often *weaker* than the original nonstandard theorem. Indeed, consider the proof of Theorem 4.7 and Remark 4.8 in which we *weakened* the antecedents of resp. (4.8) and (4.16) to resp. (4.9) and (4.17). We now obtain a version of $\mathsf{RIE}_{\mathsf{ef}}(t)$ from $\mathsf{RIE}_{\mathsf{ns}}$ from Section 4.2.2 *without* this weakening.

Theorem 4.25. *From* $\mathsf{P} \vdash \mathsf{RIE}_{\mathsf{ns}}$*, terms* i, o *can be extracted s.t.* $\textsf{E-PA}^\omega$ *proves:*

$$\begin{aligned}(\forall f,\ g, k')\Big[&(\forall k \leq i(g,k'))(\forall x, y \in [0,1])(|x-y| < \tfrac{1}{g(k)} \to |f(x)-f(y)| \leq \tfrac{1}{k})\\ &\to (\forall \pi, \pi' \in P([0,1]))\big(\|\pi\|, \|\pi'\| < \tfrac{1}{o(g,k')} \to |S_\pi(f) - S_\pi(f)| \leq \tfrac{1}{k'}\big)\Big]. \quad (4.46)\end{aligned}$$

Furthermore, $\mathsf{P} \vdash [(4.46) \to \mathsf{RIE}_{\mathsf{ns}}]$ *and from* $\textsf{E-PA}^\omega \vdash (4.46)$*, one obtains* $\mathsf{P} \vdash \mathsf{RIE}_{\mathsf{ns}}$.

Proof. The second part follows from the proof of Theorem 4.24 as follows: For standard g and k', we note that $i(g,k')$ and $o(g,k')$ are also standard. Hence, we may strengthen the antecedent of (4.46) as follows:

$$\begin{aligned}(\forall f)(\forall^{\mathrm{st}} g, k')\Big[&(\forall^{\mathrm{st}} k)(\forall x, y \in [0,1])(|x-y| < \tfrac{1}{g(k)} \to |f(x)-f(y)| \leq \tfrac{1}{k})\\ &\to (\forall \pi, \pi' \in P([0,1]))\big(\|\pi\|, \|\pi'\| < \tfrac{1}{o(g,k')} \to |S_\pi(f) - S_\pi(f)| \leq \tfrac{1}{k'}\big)\Big]. \quad (4.47)\end{aligned}$$

Note that the antecedent of (4.47) does not depend on k' and push inside this associated quantifier as follows:

$$\begin{aligned}(\forall f)(\forall^{\mathrm{st}} g)\Big[&(\forall^{\mathrm{st}} k)(\forall x, y \in [0,1])(|x-y| < \tfrac{1}{g(k)} \to |f(x)-f(y)| \leq \tfrac{1}{k})\\ &\to (\forall^{\mathrm{st}} k')(\forall \pi, \pi' \in P([0,1]))\big(\|\pi\|, \|\pi'\| < \tfrac{1}{o(g,k')} \to |S_\pi(f) - S_\pi(f)| \leq \tfrac{1}{k'}\big)\Big].\end{aligned}$$

In turn, we may strengthen '$\|\pi\|, \|\pi'\| < \frac{1}{o(g,k')}$' to '$\|\pi\|, \|\pi'\| \approx 0$' to obtain

$$\begin{aligned}(\forall f)(\forall^{\mathrm{st}} g)\Big[&(\forall^{\mathrm{st}} k)(\forall x, y \in [0,1])(|x-y| < \tfrac{1}{g(k)} \to |f(x)-f(y)| \leq \tfrac{1}{k}) \quad (4.48)\\ &\to (\forall^{\mathrm{st}} k')(\forall \pi, \pi' \in P([0,1]))\big(\|\pi\|, \|\pi'\| \approx 0 \to |S_\pi(f) - S_\pi(f)| \leq \tfrac{1}{k'}\big)\Big].\end{aligned}$$

As k' (resp. g) only appears in the final formula of the consequent (resp. the antecedent), we may push the associated quantifier inside to obtain

$$\begin{aligned}(\forall f)\Big[&(\exists^{\mathrm{st}} g)(\forall^{\mathrm{st}} k)(\forall x, y \in [0,1])(|x-y| < \tfrac{1}{g(k)} \to |f(x)-f(y)| \leq \tfrac{1}{k}) \quad (4.49)\\ &\to (\forall \pi, \pi' \in P([0,1]))\big(\|\pi\|, \|\pi'\| \approx 0 \to (\forall^{\mathrm{st}} k')(|S_\pi(f) - S_\pi(f)| \leq \tfrac{1}{k'})\big)\Big].\end{aligned}$$

The consequent of (4.49) is the definition of nonstandard integrability as in Definition 4.9 while the antecedent implies nonstandard uniform continuity (4.5). Now, the latter also implies the antecedent of (4.49), as can be seen by applying $\mathsf{HAC}_{\mathsf{int}}$ to the normal form (4.6) and taking the maximum of the resulting functional.

For the first part, recall that $\mathsf{RIE}_{\mathsf{ns}}$ implies (4.21), which yields

$$(\forall f : \mathbb{R} \to \mathbb{R})(\forall^{\mathrm{st}} g)\big[(\forall^{\mathrm{st}} k)A(f, g(k), k) \to (\forall^{\mathrm{st}} k')(\exists^{\mathrm{st}} N')B(k', N', f)\big]. \quad (4.50)$$

Bringing all standard quantifiers outside in (4.50), we obtain

$$(\forall^{\rm st} k', g)(\forall f : \mathbb{R} \to \mathbb{R})(\exists^{\rm st} N', k)\big[A(f, g(k), k) \to (\exists^{\rm st} N')B(k', N', f)\big]. \tag{4.51}$$

Applying *Idealisation* to (4.51) like in the proof of Theorem 4.11, we obtain

$$(\forall^{\rm st} k', g)(\exists^{\rm st} N', k)(\forall f : \mathbb{R} \to \mathbb{R})\big[A(f, g(k), k) \to (\exists^{\rm st} N')B(k', N', f)\big]. \tag{4.52}$$

Applying Theorem 4.2 to '$\mathsf{P} \vdash$ (4.52)' (and taking the maximum of the resulting terms as usual), one immediately obtains (4.46), as required. □

Corollary 4.26. *The system* $\mathsf{P} + (\exists^{\rm st} t)\mathsf{RIE}_{\mathsf{ef}}(t)$ *proves that a function with a* standard *modulus of uniform continuity on* $[0, 1]$ *is nonstandard integrable there.*

Proof. Repeat the first part of the proof for $(\exists^{\rm st} t)\mathsf{RIE}_{\mathsf{ef}}(t)$ instead of (4.46). □

By the previous, $\mathsf{RIE}_{\mathsf{ef}}(t)$ gives rise to a nonstandard theorem strictly *weaker* than $\mathsf{RIE}_{\mathsf{ns}}$. In particular, it is straightforward to prove that the statement *a nonstandard uniformly continuous and standard function on the unit interval has a standard modulus of uniform continuity* implies Π^0_1-TRANS. Similar results may now be obtained for *any nonstandard theorem* (like e.g. $\mathsf{MCT}_{\mathsf{ns}}$). The only difference is the omission of the *weakening* of the antecedent of (4.16) to (4.17) in Remark 4.8.

In conclusion, we have shown that the nonstandard $\mathsf{RIE}_{\mathsf{ns}}$ gives rise to the 'highly constructive' theorem (4.46). In turn, the latter gives rise to $\mathsf{RIE}_{\mathsf{ns}}$ as in Theorem 4.25. In general, *every theorem* of pure Nonstandard Analysis is 'meta-equivalent' as in Theorem 4.25 to a 'highly constructive' theorem like (4.46). Experience however bears out that this 'meta-equivalence' is usually less elegant than the one from Theorem 4.25.

4.6. **A template for the computational content of Nonstandard Analysis.** In this section, we introduce *pure* Nonstandard Analysis, i.e. that part of the latter falling within the scope of Theorem 4.2. We also formulate a template for obtaining computational content from theorems of pure Nonstandard Analysis.

Definition 4.27. [Pure Nonstandard Analysis] A theorem of *pure* Nonstandard Analysis is built up as follows.

(i) Only *nonstandard definitions* (of continuity, compactness, . . .) are used; **no** epsilon-delta definitions are used. The former have (nice) normal forms and give rise to the associated constructive definitions from Figure 2.
(ii) Normal forms are closed under *implication* by Remark 4.8 and *internal universal quantifiers* due to *Idealisation* as in (2.1).
(iii) Normal forms are closed under *prefixing a quantifier over all infinitesimals* by Theorem 4.12, i.e. if $\Phi(\varepsilon)$ has a normal form, so does $(\forall \varepsilon \approx 0)\Phi(\varepsilon)$.
(iv) Fragments of the axioms *Transfer* and *Standard Part* have normal forms.
(v) Formulas involving the *Loeb measure* (See Section 5.1) have normal forms.

We stress that item (v) should be interpreted in a specific narrow technical sense, namely as discussed in Section 5.1. We ask the reader to defer judgement until the latter section has been consulted.

The previous definition is only an approximation of all theorems within the scope of Theorem 4.2, but nonetheless covers large parts of Nonstandard Analysis. Regarding item (i) and as noted in Example 4.3, nonstandard definitions give rise to the associated constructive definition-with-a-modulus. The following list provides an overview for common notions[24], where Π_1^1-TRANS is *Transfer* limited to Π_1^1-formulas (See [81, §4]) and (μ_1) is the functional version of Π_1^1-CA_0 (See [5]).

Nonstandard Analysis definition	Constructive/functional definition
nonstandard convergence	convergence with a modulus
nonstandard continuity	continuity with a modulus
nonstandard uniform continuity	uniform continuity with a modulus
F-compactness[24]	totally boundedness
nonstandard compactness of $[0,1]$	The special fan functional Θ
nonstandard differentiability as in (5.11)	differentiability with a modulus and the derivative given
nonstandard Riemann integration	Riemann integration with a modulus
Π_1^0-TRANS	Feferman's mu-operator (μ^2)
Π_1^1-TRANS	Feferman's second mu-operator (μ_1)
STP	The special fan functional Θ

FIGURE 2. Nonstandard and constructive defitions

Finally, the following template provides a way of obtaining the computational content of theorems of pure Nonstandard Analysis.

Template 4.28. To obtain the computational content of a theorem of pure Nonstandard Analysis, perform the following steps.

(i) Convert all nonstandard definitions, formulas involving the Loeb measure, and axioms to normal forms.
(ii) Starting with the most deeply nested implication, convert implications between normal forms into normal forms using Remark 4.8.
(ii′) If 'meta-equivalence' as in Section 4.5 is desired, omit the weakening from (4.15) to (4.17) in Remark 4.8.
(iii) When encountering quantifiers over all infinitesimals, use Theorem 4.12 to obtain a normal form. When encountering internal universal quantifiers, apply *Idealisation* as in (2.1).
(iv) When a normal form has been obtained, apply Theorem 4.2 and modify the resulting term as necessary, usually by taking the maximum of the finite sequence at hand.

We have now completed the main part of this paper, namely to establish the basic claims made in the introduction. For the rest of this paper, we will consider related *more advanced* results.

We finish this section with the observation by Hao Wang that constructive mathematics may be viewed as a "mathematics of doing" while classical mathematics is a "mathematics of being" (See e.g. [73, Preface]). Nonstandard Analysis then fits into this view as follows: In the second chapter of the *Tao Te Ching*, the philosopher Lao Tse writes *the sage acts by doing nothing*, and one could view the associated notion of *wu wei* as the founding principle of Nonstandard Analysis.

[24]A space is *F-compact* in NSA if there is a discrete grid which approximates every point of the space up to infinitesimal error, i.e. the intuitive notion of compactness from physics and engineering (See [81, §4]).

5. A grab-bag of Nonstandard Analysis

We discuss various advanced results regarding Nonstandard Analysis and its (local) constructivity.

5.1. **The computational content of the Loeb measure.** The *Loeb measure* ([45,107]) is one of the cornerstones of *Nonstandard Analysis* providing an elegant and universal approach to measure theory. The traditional development of the Loeb measure makes use of the non-constructive axiom *Saturation* and *external sets* (not present in IST pur sang). In this section, we discuss the possibility of formulating the Loeb measure in P, and this measure's potential computational content.

As to prior art, a special case of the Loeb measure is introduced in a weak fragment of Robinsonian Nonstandard Analysis in [88] *using external sets*, but without the use of *Saturation*; the definition of *measure* from Reverse Mathematics (See [87, X.1]) is used. On the internal side, the system P extended with the (countable) axiom *Saturation* CSAT can be given computational meaning; in particular Theorem 4.2 holds for the extension $\mathsf{P}+\mathsf{CSAT}$ (See [8,31]), but t then involves *bar recursion*, an advanced computation scheme which (somehow) embodies the computational content of the axiom of choice ([54]).

Now, we have established in the previous sections that *pure Nonstandard Analysis* contains plenty of computational content *not involving bar recursion*; We believe that the study of the Loeb measure should be *similarly elementary*, *lest history repeat itself* in the form of incorrect claims that Nonstandard Analysis is somehow non-constructive ([82]), this time around based on the use of bar recursion in the study of the Loeb measure. Serendipitously, the author has undertaken first steps towards such elementary study of the Loeb measure in [84], as discussed now.

At first sight, mainstream measure theory seems hopelessly non-constructive: we know from Reverse Mathematics that the existence of the Lebesgue measure $\lambda(A)$ for open sets A already gives rise to the Turing jump ([87, X.1]). Indeed, the definition of λ on $[0,1]$ in Reverse Mathematics is as follows:

Definition 5.1. [Lebesgue measure λ] For $\|g\|:=\int_0^1 g(x)dx$, $U\subseteq[0,1]$,

$$\lambda(U):=\sup\{\|g\|: g\in C([0,1])\wedge 0\le g\le 1\wedge(\forall x\in[0,1]\setminus U)(g(x)=0)\}.$$

The existence of this supremum cannot be proved in computable mathematics, like the base theory RCA_0 of Reverse Mathematics from [87, I], or constructive mathematics. However, there is a way around this 'non-existence problem' pioneered in [110,111]: the formulas $\lambda(U)>_{\mathbb{R}}0$ and $\lambda(U)=_{\mathbb{R}}0$ *make perfect*[25] *sense*, even if the supremum in Definition 5.1 does not exist. Furthermore, most theorems of measure theory only involve the Lebesgue measure (and other measures) via such formulas. In other words, *even if the Lebesgue measure cannot be defined in general*, we can still write down most theorems of measure theory without any problems, even in the language of second-order arithmetic typical of Reverse Mathematics.

Now, the *Loeb measure* L is defined as a *nonstandard* supremum or infinimum (See e.g. [45,108]). Similar to the Lebesgue measure, this supremum does not necessarily exist in weak systems of Robinsonian Nonstandard Analysis. However,

[25] In particular, $\lambda(U)>_{\mathbb{R}}0$ is $(\exists g\in C([0,1]))(\|g\|>_{\mathbb{R}}0\wedge 0\le g\le 1\wedge(\forall x\in[0,1]\setminus U)(g(x)=0))$, and similar for $\lambda(U)=_{\mathbb{R}}0$.

formulas like '$L(A) \gg 0$' or '$L(A) \approx 0$' make perfect sense in such weak systems, in the same way as for the Lebesgue measure, as explored in [88]. Hence, most theorems involving the Loeb measure may be studied in this 'Reverse Mathematics' way, even though the Loeb measure does not necessarily exist in general. Note that *external sets* are used in an essential way.

Finally, inspired by the results in [88], the novel features introduced in [84] regarding the Loeb measure on $[0, 1]$ are as follows:

(i) The formulas '$L(A) \gg 0$' or '$L(A) \approx 0$' make perfect sense in P, i.e. can be formulated in the language of P *without the use of external sets*.
(ii) If the set A is given by a normal form, then the formulas '$L(A) \gg 0$' or '$L(A) \approx 0$' also have a normal form.

The first item implies that the *external* formulation of the Loeb measure (using the standard part map) from [88] can be expressed as a formula in the language of P. The second item means that if $(\forall z \in \mathbb{R})(z \in A \leftrightarrow (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\varphi(z, y, z))$ where φ is internal, then the formulas '$L(A) \gg 0$' or '$L(A) \approx 0$' also have a normal form in P.

In conclusion, we can study most theorems regarding the Loeb measure in the system P, in much the same way as is done in Reverse Mathematics, by item (i). Furthermore, these theorems have lots of computational content by item (ii), in light of the generality of the class of normal forms as discussed in Section 4.6.

5.2. **Constructive features of Nonstandard Analysis.** In this section, we discuss some constructive features of Nonstandard Analysis.

5.2.1. *Metamathematical properties.* While there is no official formal definition of what makes a logical system 'constructive', there are a number of metamathematical properties which are often considered to be hallmarks of intuitionistic/constructive theories. Rathjen explores these properties in [71] for *Constructive Zermelo-Fraenkel Set Theory* CZF, one of the main foundational systems for constructive mathematics. We discuss to what extent the system P and its extensions entertain nonstandard versions of the nine properties listed in [71, §1].

First of all, a system S is said to have the *existence property* **EP** if it satisfies the following: Let $\varphi(x)$ be a formula with at most one free variable x; if $S \vdash (\exists x)\varphi(x)$ then there is some formula $\theta(x)$ with only x free such that $S \vdash (\exists! x)(\theta(x) \wedge \varphi(x))$, where the '!' denotes unique existence. Now, P has the following *nonstandard existence property*: If $\mathsf{P} \vdash (\exists^{\mathrm{st}} x^\rho)\varphi(x)$ with $\varphi(x)$ internal and at most one free variable x, then there is an internal formula $\theta(x)$ such that $\mathsf{P} \vdash (\exists^{\mathrm{st}}! x^\rho)(\theta(x) \wedge \varphi(x))$. Indeed, applying Theorem 4.2 to '$\mathsf{P} \vdash (\exists^{\mathrm{st}} x^\rho)\varphi(x)$', we obtain a term t^{ρ^*} such that $\mathsf{E\text{-}PA}^\omega \vdash (\exists x^\rho \in t)\varphi(x)$ and can define $\theta(x) := (\exists j < |t|)(x =_\rho t(j) \wedge (\forall i < j)\neg\varphi(t(i)))$. Since t (and $|t|$ and $t(j)$ for $j < |t|$) is standard in P by the basic axioms from Definition A.3, $\theta(x)$ implies that $\mathrm{st}(x)$. The previous also holds for any internal extension of P, as for the below properties we consider now.

Secondly, a system S is said to have the *numerical existence property* **NEP** if it satisfies the following: Let $\varphi(x)$ be a formula with at most one free variable x; if $S \vdash (\exists n^0)\varphi(n)$ then there is some numeral $\overline{n}$ such that $S \vdash \varphi(\overline{n})$. Now, P has the following *nonstandard numerical existence property*: If $\mathsf{P} \vdash (\exists^{\mathrm{st}} x^0)\varphi(x)$ with $\varphi(x)$ internal and at most one free variable x, then there is a numeral $\overline{n}$ such that $\mathsf{P} \vdash (\exists x^0 \leq \overline{n})\varphi(x)$. The latter is proved in the same way as the previous property.

Thirdly, a system S is said *closed under Church's rule* **CR** if it satisfies the following: Let $\varphi(x, y)$ be a formula with at most the free variables shown; if $S \vdash (\forall m^0)(\exists n^0)\varphi(n)$ then there is some numeral $\overline{e}$ such that $S \vdash (\forall m^0)\varphi(m, \{\overline{e}\}(m))$ where $\{e\}(n)$ is the value of the e-th Turing machine with input n. Now, P is closed under *nonstandard Church's rule*: If $\mathsf{P} \vdash (\forall^{\text{st}} m^0)(\exists^{\text{st}} n^0)\varphi(m, n)$ with $\varphi(x, y)$ internal and at most the free variable shown, then there is a numeral $\overline{e}$ such that $\mathsf{P} \vdash (\forall m^0)(\exists n^0 \leq \{\overline{e}\}(n))\varphi(m, n)$. The latter is proved in the same way as the previous property; the term produced in this way is ε_0-computable as it is a term definable in Peano arithmetic. A similar nonstandard version of *the first variant of Church's rule* $\mathbf{CR}_1$ and *Extended Church's rule* **ECR** can be obtained, assuming the antecedent in the latter has the form $(\forall^{\text{st}} y)\varphi(y, x^0)$ for internal φ.

On a related note, P and H prove the following nonstandard version of Church's thesis: $(\forall^{\text{st}} f^1)(\exists e^0)(\forall^{\text{st}} x^0)(f(x) = \varphi_e(x))$. This can be proved by applying *Idealisation* to the trivial formula $(\forall^{\text{st}} f^1, x^0)(\exists e^0)(f(x) = \varphi_e(x))$, or simply noting that there are N^N different (primitive recursive) functions mapping $\{0, 1, \ldots, N-1\}$ to $\{0, 1, \ldots, N-1\}$ (and zero otherwise). Hence, nonstandard Church's thesis merely expresses that *from the point of view of the nonstandard universe* all standard functions are computable when limited to the standard universe.

Fourth, since disjunction '$\vee$' gives rise to internal formulas, this symbol has no computational content (in contrast to '$(\exists^{\text{st}})$'). As a result the *Disjunction property* **DP** and *Unzerlegbarkeits rule* **UZR** do not apply to P. The *Uniformity rule* **UR** is similar to idealisation, but only with essential modification.

In conclusion, we again emphasise that we do **not** claim that P is somehow part of constructive mathematics; we merely point out that the standard universe of P satisfies versions of the typical properties of logical systems of constructive mathematics. In other words, we again observe that P is 'too constructive' to be called non-constructive (but does include LEM and is hence not constructive either).

5.2.2. *The axiom of extensionality.* We discuss the constructive status of the axiom of extensionality, and how this axiom relates to our system P.

First of all, Martin-Löf has proposed his *intuitionistic type theory* as a foundation for (constructive) mathematics ([60]). These systems are close to computer programming and form the basis for various proof assistants, including Coq, Agda, and Nuprl ([106]). Martin Löf's early type theories were *extensional* (as is Nurpl), but the later systems were *intensional* (as is Agda). An advantage of intensional mathematics is the decidability of type checking, while a disadvantage is that every (even very obvious) equality has to be proved separately. The practice of Nuprl suggests however that checking whether an expression is correctly typed, is usually straightforward to perform by hand *in practice*, i.e. decidable type checking is not a *conditio sine qua non* when formalising mathematics.

Secondly, the system P of Nonstandard Analysis obviously includes the *internal* axiom of extensionality. However, internal axioms are ignored by the term extraction algorithm in Theorem 4.2. In particular, computational content is extracted from certain statements about the *standard universe*, and it is a natural question whether the latter satisfies the 'standard' axiom of extensionality $(\mathsf{E})^{\text{st}}$ as follows:

$$(\forall^{\text{st}} x^\rho, y^\rho, \varphi^{\rho \to \tau})\big[x \approx_\rho y \to \varphi(x) \approx_\tau \varphi(y)\big].$$

where the notations are taken from Remark A.12. Clearly, the previous sentence follows immediately from the (internal) axiom of extensionality by *Transfer*, which is however absent in P. In particular, the standard universe in P is actually highly *intensional* in the sense of the following theorem.

Theorem 5.2. *Let Δ_{int} be any collection of internal formulas such that $\mathsf{P} + \Delta_{\mathsf{int}}$ is consistent. Then the latter cannot prove the following:*

$$(\forall^{\mathrm{st}} F^2, f^1, g^1)(f \approx_1 g \to F(f) =_0 F(g)) \tag{5.1}$$

Proof. Since $f \approx_1 g$ is just $(\forall^{\mathrm{st}} N)(\overline{f}N =_0 \overline{g}N)$, (5.1) implies the normal form

$$(\forall^{\mathrm{st}} F^2, f^1, g^1)(\exists^{\mathrm{st}} N)(\overline{f}N =_0 \overline{g}N \to F(f) =_0 F(g)). \tag{5.2}$$

A proof of (5.2) in $\mathsf{P} + \Delta_{\mathsf{int}}$ provides a term t from Gödel's T which realises the axiom of extensionality for type two functionals (over $\mathsf{E\text{-}PA}^\omega + \Delta_{\mathsf{int}}$) as follows:

$$(\forall F^2, f^1, g^1)(\exists N \leq t(F, f, g))(\overline{f}N =_0 \overline{g}N \to F(f) =_0 F(g)). \tag{5.3}$$

However, Howard has shown in [44] that a term as in (5.3) does not exist. □

Howard's results in [44] also imply that standard extensionality as in (5.1) for type three functionals cannot be proved in any extension of P which is conservative over *Zermelo-Fraenkel set theory* ZF. Furthermore, we can neatly classify the strength of (5.1) as follows, where $\mathsf{TJ}(\varphi, f)$ is $(\exists^2)$ from Section 4.3 without the two outermost quantifiers.

Theorem 5.3. *The system* P *proves* $\Pi^0_1\text{-}\mathsf{TRANS} \leftrightarrow [(5.1) \wedge (\exists^{\mathrm{st}}\varphi^2)(\forall^{\mathrm{st}} f^1)\mathsf{TJ}(\varphi, f)]$.

Proof. For the forward implication, $\Pi^0_1\text{-}\mathsf{TRANS}$ applied to the internal axiom of extensionality for standard F^2, f^1, g^1 immediately yields (5.2). Also, $\Pi^0_1\text{-}\mathsf{TRANS}$ implies (4.34), and latter yields the second conjunct in the right-hand side of the equivalence. For the reverse implication, suppose $\Pi^0_1\text{-}\mathsf{TRANS}$ is false, i.e. there is standard h^1 such that $(\forall^{\mathrm{st}} n^0)(h(n) = 1) \wedge (\exists m^0)(h(m) = 0)$. Clearly, $h \approx_1 11\ldots$ while for φ as in $(\forall^{\mathrm{st}} f^1)\mathsf{TJ}(\varphi, f)$ we have $\varphi(h) = 0 \neq \varphi(11\ldots)$, contradicting extensionallity as in (5.1), and $\Pi^0_1\text{-}\mathsf{TRANS}$ follows. □

In conclusion, the standard universe of P is highly intensional in the sense of the previous theorem. Thus, if one believes that intensionality is essential/important/... for the constructive nature of a logical system, then the above results provide a partial explanation why P has so much computational content.

5.2.3. *Decomposing the continuum.* In this section, we discuss the *indecomposable* nature of the *continuum* in intuitionism. We will show that the continuum in Nonstandard Analysis is **in**decomposable *from the point of view of the standard universe*, and decomposable *from the point of view of the nonstandard universe.*

The *continuum*, i.e. the set of real numbers $\mathbb{R}$, is *indecomposable* in intuitionistic mathematics, which means that if $\mathbb{R} = A \cup B$ and $A \cap B = \varnothing$, then $A = \mathbb{R}$ or $B = \mathbb{R}$. Brouwer first published this result in [21] (See also [101, p. 490]) and a modern treatment by van Dalen may be found in [100]. This in-decomposability result of course hinges on the absence of the law of excluded middle in intuitionistic mathematics, as e.g. $(\forall x \in \mathbb{R})(x > 0 \vee x \leq 0)$ allows us to split $\mathbb{R}$ as $\{x \in \mathbb{R} : x > 0\} \cup \{x \in \mathbb{R} : x \leq 0\}$. An intuitive description of the above is as follows:

> In intuitionistic mathematics the situation is different; the continuum has, as it were, a syrupy nature, one cannot simply take away one point. In the classical continuum one can, thanks to the principle of the excluded third, do so. To put it picturesquely, the classical continuum is the frozen intuitionistic continuum. If one removes one point from the intuitionistic continuum, there still are all those points for which it is unknown whether or not they belong to the remaining part. ([100, p. 1147])

We now show that the continuum in Nonstandard Analysis is indecomposable in one sense, and decomposable in another sense. To this end, let $N \in \mathbb{N}$ be nonstandard and let $[x](N)$ be the N-th rational approximation of x. Since inequality on $\mathbb{Q}$ is decidable (as opposed to the situation for '$<_{\mathbb{R}}$'), we have

$$\mathbb{R} = \{x \in \mathbb{R} : [x](N) >_{\mathbb{Q}} 0\} \cup \{x \in \mathbb{R} : [x](N) \leq_{\mathbb{Q}} 0\}, \tag{D}$$

where the first set is denoted A_N and the second one B_N. While (D) provides a (decidable) decomposition of $\mathbb{R}$, there are problems: the decomposition is *fundamentally arbitrary* and based on an *nonstandard* algorithm in an essential way.

Firstly, as to the arbitrariness of (D), it is intuitively clear that for $x \approx 0$ we can have $x \in A_N$ but $x \in B_M$ for nonstandard $N \neq M$, i.e. the decomposition (D) seems *non-canonical* or *arbitrary* as it depends on the choice of nonstandard number. To see that this arbitrariness is fundamental, consider (5.4) which expresses that (D) is *not* arbitrary in the aforementioned sense for standard reals:

$$(\forall^{\text{st}} x \in \mathbb{R})(\forall N, M \in \mathbb{N})\big((\neg\text{st}(N) \wedge \neg\text{st}(M)) \to x \in A_N \leftrightarrow x \in A_M\big). \tag{5.4}$$

By the following theorem, (5.4) is *Transfer* in disguise, and hence fundamentally non-constructive by Corollary 4.17. As a consequence, the decomposition (D) is always fundamentally arbitrary in any system where the Turing jump as in (μ^2) is absent, like e.g. in constructive mathematics or P.

Theorem 5.4. *The system* P *proves* Π^0_1-$\mathsf{TRANS} \leftrightarrow (5.4)$.

Proof. For the forward direction, note that for a standard real x, $[x](N) > 0$ implies $(\exists^{\text{st}} n)([x](n) > 0)$ by *Transfer*, and (5.4) is immediate. For the reverse direction, suppose $\neg\Pi^0_1$-TRANS, i.e. there is standard h^1 such that $(\forall^{\text{st}} n^0)(h(n) = 0) \wedge (\exists m^0)(h(m) \neq 0)$, and define the real x_0 as: $[x_0](k) = 0$ if $(\forall i \leq k)(h(i) = 0)$, and $\frac{1}{2^{i_0}}$ if i_0 is the least $i \leq k$ such that $h(i) \neq 0$. Clearly, $x \in B_N$ for small enough nonstandard N, and $x \in A_N$ for large enough nonstandard N. This contradiction with (5.4) yields the reverse direction. $\square$

The real x_0 defined in the proof of the theorem is an example of the aforementioned 'syrupy' nature of the continuum: this real (and many like it) can be zero or positive in different models of P, depending whether *Transfer* holds.

Secondly, as to the nonstandardness of (D), there clearly exists F^2 with

$$(\forall x \in \mathbb{R})(F(x) = 0 \to x \in A_N \wedge F(x) = 1 \to x \in B_N), \tag{5.5}$$

and F is even given by an algorithm (involving nonstandard numbers). However, by Theorem 4.2, we can only obtain computational information from *standard objects*, and it is thus a natural question if there is *standard* F as in (5.5). Now, from the existence of such F, one readily derives WKL^{st}, i.e. *weak König's lemma* ([87, IV]) relative to 'st', which is fundamentally non-constructive and cannot be derived in

P. The same derivation goes through for A_N, B_N replaced by *any* internal A, B. Thus, while (D) constitutes a *decidable* decomposition of $\mathbb{R}$, the associated decision procedure is *fundamentally nonstandard*.

In conclusion, while $\mathbb{R}$ has a decidable decomposition (D) in the system P, the decomposition is *fundamentally arbitrary* and based on an *nonstandard* algorithm in an essential way. If one requires a decomposition be *canonical* or to have *computational content* (and therefore be standard), then (D) is disqualified.

Let us rephrase this situation as follows: the continuum in Nonstandard Analysis is indecomposable *from the point of view of the standard universe*, and decomposable *from the point of view of the nonstandard universe.* Since we can only extract computational information from the standard universe, our results are in line with intuitionistic mathematics.

5.2.4. *Tennenbaum's theorem.* We discuss the (non-)constructive nature of Nonstandard Analysis in light of *Tennenbaum's theorem* (See e.g. [51, §11.3]).

As noted in Section 3.5, even fragments of Robinson's Nonstandard Analysis based on arithmetic seem fundamentally non-constructive. Indeed, *Tennenbaum's theorem* as formulated in [51, §11.3] 'literally' states that any nonstandard model of Peano Arithmetic is not computable.

What is meant is that for a nonstandard model $\mathcal{M}$ of Peano Arithmetic, the operations $+_{\mathcal{M}}$ and $\times_{\mathcal{M}}$ cannot be computably defined in terms of the operations $+_{\mathbb{N}}$ and $\times_{\mathbb{N}}$ of the standard model $\mathbb{N}$ of Peano Arithmetic. Similar results exist for fragments ([51, §11.8]) and Nonstandard Analysis thus *seems* fundamentally non-constructive even at the level of basic arithmetic.

Now, while certain nonstandard models indeed require non-constructive tools to build, models are not part of Nelson's *axiomatic approach* to Nonstandard Analysis via IST. What is more, the latter explicitly disallows the formation of external sets like '*the operation + restricted to the standard numbers*'. Nelson specifically calls attention to this rule on the first page of [62] introducing IST:

> *We may not use external predicates to define subsets.* We call the violation of this rule illegal set formation. (Emphasis in original)

To be absolutely clear, one of the fundamental components of Tennenbaum's theorem, namely the *external set* '+ restricted to the standard naturals' is missing from *internal* set theory IST, as the latter exclusively deals with *internal* sets. Thus, we may claim that Tennenbaum's theorem is merely an artefact of the model-theoretic approach to Nonstandard Analysis.

Finally, Connes' critique of Nonstandard Analysis mentioned in Section 3.5 seems to be based on similarly incorrect observations, namely that the models generally used in Robinsonian Nonstandard Analysis are fundamentally non-constructive and *therefore* Nonstandard Analysis is too. By way of an analogy: non-constructive mathematics (including models) is used in physics all the time, but does that imply that physical reality is therefore non-constructive (in the sense that there exist non-computable objects out there)?

5.2.5. *Bishop's numerical implication.* We discuss the, as it turns out intimate, connection between Bishop's notion of *numerical implication* from [15] and some of the axioms of H from Definition A.7.

Bishop introduces a new connective *numerical implication* (aka Gödel implication) in [15, p. 60] as an alternative to the implication present in intuitionistic logic with the associated BHK interpretation, motivated as follows:

> The most urgent foundational problem of constructive mathematics concerns the numerical meaning of implication. ([15, p. 56])

Numerical implication is based on Gödel's Dialectica interpretation ([37]) and is introduced because Bishop deems the numerical meaning of the BHK-implication to be unclear in general.

Bishop points out that numerical implication and BHK-implication amount to the same thing *in practice*, and even derives the former from the latter, however using two *non-constructive* steps (See [15, p. 56-60]). In particular, [15, (4) $\to$ (5)] is an instance of the *independence of premises principle*, while [15, (7) $\to$ (12)] is a generalisation of *Markov's principle*.

Now, the term extraction property of P and H is based on a nonstandard version of the Dialectica interpretation, called the *nonstandard Dialectica interpretation* D_{st} (See [9, §5]). Furthermore, the system H includes (See Definition A.7) the axiom $\mathsf{HIP}_{\forall^{\text{st}}}$, a version of the *independence of premises principle*, and the axiom $\mathsf{HGMP}^{\text{st}}$, a generalisation of *Markov's principle*. By the following theorem, the latter two principles are exactly what is needed to show (inside H) that the class of normal forms is closed under implication, similar to Bishop's derivation of numerical implication from BHK-implication in [15].

Theorem 5.5. *The system* H *proves that a normal form can be derived from an implication between normal forms.*

Proof. Similar to the derivation in Remark 4.8, we work in H and consider

$$(\forall^{\text{st}} x)(\exists^{\text{st}} y)\varphi(x, y) \to (\forall^{\text{st}} z)(\exists^{\text{st}} w)\psi(z, w), \tag{5.6}$$

where φ, ψ are internal. Since standard functionals yield standard output from standard input, (5.6) implies

$$(\forall^{\text{st}} \zeta)\big[(\forall^{\text{st}} x)\varphi(x, \zeta(x)) \to (\forall^{\text{st}} z)(\exists^{\text{st}} w)\psi(z, w)\big]. \tag{5.7}$$

We can bring outside the quantifier over z (even in H) as follows:

$$(\forall^{\text{st}} \zeta, z)\big[(\forall^{\text{st}} x)\varphi(x, \zeta(x)) \to (\exists^{\text{st}} w)\psi(z, w)\big], \tag{5.8}$$

and the formula in square brackets turns out to have exactly the right form for $\mathsf{HIP}_{\forall^{\text{st}}}$. Applying the latter yields

$$(\forall^{\text{st}} \zeta, z)(\exists^{\text{st}} W)\big[(\forall^{\text{st}} x)\varphi(x, \zeta(x)) \to (\exists w \in W)\psi(z, w)\big], \tag{5.9}$$

which in turn has exactly the right form to apply $\mathsf{HGMP}^{\text{st}}$, yielding

$$(\forall^{\text{st}} \zeta, z)(\exists^{\text{st}} V, W)\big[(\forall x \in V)\varphi(x, \zeta(x)) \to (\exists w \in W)\psi(z, w)\big], \tag{5.10}$$

which is a normal form since the formula in square brackets in (5.10) is internal. $\square$

In conclusion, the previous points to similarities between Bishop's numerical implication and Nonstandard Analysis, but we do not have any deeper insights beyond the observed analogies.

5.2.6. *The meaning of Nonstandard Analysis.* In this section, we discuss a possible interpretation of the predicate 'x is standard' from Nonstandard Analysis. In particular, we point out similarities between this predicate and predicates pertaining to computational efficiency in constructive mathematics.

First of all, set theory for instance typically does **not** deal with questions like 'What is a set?' or the meaning of the symbol '$\in$' for elementhood. These questions do come up, especially in the philosophy of mathematics, but are mostly absent from the mathematics itself. The same holds for the 'st' predicate in Nonstandard Analysis (and e.g. infinitesimals), as summarised by Nelson as follows:

> In addition to the usual undefined binary predicate $\in$ of set theory we adjoin a new undefined unary predicate *standard*. [...] To assert that x is a standard set has no meaning within conventional mathematics-it is a new undefined notion. ([62, p. 1165])

Nonetheless, experience bears out that many newcomers to Nonstandard Analysis do ask questions regarding the meaning of 'st' and infinitesimals. While the author mostly agrees with Nelson's quote, we did formulate an intuitive way of understanding infinitesimals in Section 4.1, namely as an elegant shorthand for computational content like moduli. Paying homage to the imaginary proverb 'in for an infinitesimal, in for the entire framework', we now accommodate the aforementioned newcomers and indulge in the quest for the meaning of the new predicate 'st'.

Secondly, by the nature of the BHK-interpretation, all connectives have computational content in constructive mathematics. However, for a given theorem stating the existence of x, it is possible one does not need *all* this computational content to compute x, i.e. some of this computational content is superfluous. By way of an example, one *only* needs a modulus of uniform continuity of f to compute the modulus of integration of f in $\mathsf{RIE}_{\mathsf{ef}}(t)$ from Section 4.2.2, i.e. one does not need f itself. Hence, if one is interested in *efficient computation*, it makes sense to ignore all superfluous computational content, and develop a mechanism which can perform this task (by hand or automatically).

Now, the previous considerations are the motivation for Berger's *Uniform Heyting arithmetic* ([10]) in which quantifiers '$\exists^{\mathrm{c}}$' and '$\exists^{\mathrm{nc}}$' are introduced which may be read as 'computationally relevant' and 'computationally irrelevant' existence. The computational content of connectives decorated with 'nc' is ignored, leading to more efficient algorithms. The proof-assistant *Minlog* ([89]) even sports a *decoration algorithm* to automatically decorate a proof with the 'c' and 'nc' predicates. Furthermore, the proof-assistant *Agda* ([1]) has the 'dot notation' which has the same functionality as the 'nc' predicate, while *homotopy type theory* ([112]) includes 'truncated existence' $\|\Sigma\|$ also similar to '$\exists^{\mathrm{nc}}$'. Finally, aspects of the previous may already be found in the work of Lifschitz ([46]).

As discussed in [81], [9, p. 1963], and [38, §4], the predicate 'st' in P seems to behave very similarly to the predicate 'c' for 'computational relevance', while the role of '$\forall^{\mathrm{nc}}$' from Minlog is played by '$\forall$' in Nonstandard Analysis. By way of an example, consider $\mathsf{RIE}_{\mathsf{ns}}$ from Section 4.2.2 and observe that in the latter the quantifier '$(\forall f)$' is present, rather than '$(\forall^{\mathrm{st}} f)$'. As a result, the extracted term t in $\mathsf{RIE}_{\mathsf{ef}}(t)$ does **not** depend on f, i.e. '$(\forall f)$' is not computationally relevant, while $(\forall^{\mathrm{st}} f)$ would be. Furthermore, the connectives '$\vee$' and '$\rightarrow$' retain their usual classical/non-constructive behaviour in P, which corresponds to $\vee^{\mathrm{nc}}$ and $\rightarrow^{\mathrm{nc}}$ in

Minlog. Thus, we obtain an **informal** analogy between constructive mathematics and classical Nonstandard Analysis (without *Transfer*) as summarised in the following table. We cannot stress the *heuristic* nature of this comparison enough.

Homotopy Type Theory	Minlog / Uniform HA	Nonstandard Analysis as in system P
Σ	$\exists^{\mathrm{c}}$	$\exists^{\mathrm{st}}$
Π	$\forall^{\mathrm{c}}$	$\forall^{\mathrm{st}}$
$\|\Sigma\|$	$\exists^{\mathrm{nc}}$	$\exists$
	$\forall^{\mathrm{nc}}$	$\forall$
	$\vee^{\mathrm{nc}}$	$\vee$
	$\rightarrow^{\mathrm{nc}}$	$\rightarrow$

We also stress that we do not claim P to be a constructive system, but we also point out that P does sport too many of properties of constructive mathematics to be dismissed as non-constructive. Therefore, it occupies the twilight zone between the constructive and non-constructive.

5.2.7. *Efficient algorithms and Nonstandard Analysis.* While most of this paper is foundational in nature, we now argue that Theorem 4.2 holds the promise of providing *efficient* terms via term extraction applied to formalised proofs. The implementation in the proof assistant Agda of the term extraction algorithm of Theorem 4.2 is underway by the author and Chuangjie Xu in [109].

First of all, Nonstandard Analysis generally involves *very short* proofs (compared to proofs in mainstream mathematics). Short proofs are one heuristic (among others) for extracting *efficient* terms in *proof mining* ([56]).

Secondly, as noted in Theorem 4.2, *internal axioms do not contribute to the extracted term.* In particular, the extracted terms are determined *only* by the external axioms of P; *all axioms of 'usual' mathematics, including mathematical induction, do not influence the extracted term.* Note that *external induction* $\mathsf{IA}^{\mathrm{st}}$ (See Definition A.3) is rarely used in the practice of Nonstandard Analysis.

Thirdly, as can be gleaned from the proof of [9, Theorem 5.5], terms extracted from external axioms of H and P *other than external induction* are of trivial complexity, except for **one** relating to the $(\forall^{\mathrm{st}})$-quantifier (See [9, p. 1982]) which involves sequence concatenation. Furthermore, one rarely uses external induction in practice while certain[26] 'basic' applications of external induction yield low-complexity terms too.

In conclusion, the three previous observations hold the promise of providing *efficient* terms via term extraction applied to formalised proofs as in Theorem 4.2.

[26]As explored in [109], applying Theorem 4.1 to *if* $m \leq_0 n$ *for standard* n*, then* m *is standard* yields a function $g^{0\rightarrow 0^*}$ such that $g(n) = \langle 0, 1, \ldots, n\rangle$. The latter function is standard as the recursor constants and 0 and + are, and thus yields a proof of the original nonstandard statement.

5.3. **Constructive Nonstandard Analysis versus local constructivity.** While most of this paper deals with *classical* Nonstandard Analysis, we now discuss *constructive* Nonstandard Analysis and its connection to local constructivity. We already have the system H as an example of the syntactic approach to constructive Nonstandard Analysis, and we now discuss the *semantic approach*.

Intuitively, the *semantic* approach to Nonstandard Analysis pioneered by Robinson ([74]) consists in somehow building a nonstandard model of a given structure (say the set of real numbers $\mathbb{R}$) and proving that the original structure is a strict subset of the nonstandard model (usually called the set of *hyperreal numbers* ${}^*\mathbb{R}$) while establishing properties similar to *Transfer*, *Idealisation* and *Standardisation* as theorems of this model and the original structure. Historically, Nelson of course studied Robinson's work and axiomatised the semantic approach in his internal set theory IST. The most common way of building a suitable nonstandard model is using a *free ultrafilter* (See e.g. [45, 108]). The existence of the latter is a rather strong *non-constructive* assumption, slight weaker than the axiom of choice of ZFC.

As it turns out, building nonstandard models with properties like *Transfer* can also be done constructively: Palmgren in [70, Section 2] and [69] constructs a nonstandard model $\mathcal{M}$ (also called a 'sheaf' model) satisfying the *Extended Transfer Principle* by [70, Corollary 4 and Theorem 5] (See also [61]). As noted by Palmgren ([70, p. 235]), the construction of $\mathcal{M}$ can be formalised in Martin-Löf's constructive type theory ([60]). The latter was developed independently of Bishop's constructive mathematics, but can be viewed as a foundation of the latter.

It should be clear by now that there is a fundamental difference between our approach and that of Palmgren and Moerdijk: The latter attempt to mimic Robinson's approach as much as is possible inside a framework for constructive mathematics, while we attempt to bring out the (apparently copious) computational content of *classical Nonstandard Analysis itself* using the system P and its extensions. The constructivist Stolzenberg has qualified the former 'mimicking' approach as *parasitic* (or *scavenger*) as follows, though we do not share his view.

> Nowadays, what is called "constructive" mathematical practice consists in taking a classical theorem or theory that is a product of ordinary classical mathematical practice and trying to produce a "good" constructive counterpart to it. Obviously, this enterprise is parasitic on the theorems and theories of ordinary classical practice. ([92])

We only mention Stolzenberg's view as a means of launching a discussion concerning which place our results on Nonstandard Analysis have in the (apparently rather emotionally charged) constructive pantheon.

Finally, we establish that our approach cannot easily be reconciled with the approach by Palmgren and Moerdijk. To this end, we show that principles (built into the Palmgren-Moerdijk framework) connecting epsilon-delta and nonstandard definitions give rise to non-constructive oracles when considered in our framework. To this end, let NSD be the statement *a standard* $f : \mathbb{R} \to \mathbb{R}$ *differentiable at zero is also* nonstandard differentiable *there*, where the latter is as follows.

Definition 5.6. A function f is *nonstandard differentiable* at a if

$$(\forall \varepsilon, \varepsilon' \neq 0)\big(\varepsilon, \varepsilon' \approx 0 \to \tfrac{f(a+\varepsilon)-f(a)}{\varepsilon} \approx \tfrac{f(a+\varepsilon')-f(a)}{\varepsilon'}\big), \tag{5.11}$$

Now, NSD is a theorem of IST but we also have the following implication.

Theorem 5.7. *The system* $\mathsf{P} + \mathsf{NSD}$ *proves* Π^0_1-TRANS.

Proof. Working in $\mathsf{P}+\mathsf{NSD}$, suppose $\neg\Pi^0_1$-TRANS, i.e. there is standard h^1_0 such that $(\forall^{\text{st}}n)h_0(n) = 0$ and $(\exists m_0)h(m_0) \neq 0$. Define the standard real x_0 as $\sum_{n=0}^{\infty} \frac{h(n)}{2^n}$. Since $0 \approx x_0 >_{\mathbb{R}} 0$ the standard function $f_0(x) := \sin(\frac{x}{x_0})$ is clearly well-defined and differentiable in the usual internal 'epsilon-delta' sense. However,

$$\tfrac{f_0(x_0)-f_0(0)}{x_0} =_{\mathbb{R}} \tfrac{\sin(1)}{x_0} \not\approx \tfrac{2}{\pi x_0} =_{\mathbb{R}} \tfrac{f_0(\frac{\pi}{2}x_0)-f_0(0)}{\frac{\pi}{2}x_0}$$

which implies that f_0 is not nonstandard differentiable at zero. This contradiction yields Π^0_1-TRANS, and we are done. $\square$

Corollary 5.8. *From the proof that* $\mathsf{P} \vdash \mathsf{NSD} \to \Pi^0_1$-$\mathsf{TRANS}$*, a term t can be extracted such that* $\mathsf{E\text{-}PA}^\omega \vdash (\forall \Xi^3)\big(\mathsf{DIF}(\Xi) \to \mathsf{MU}(t(\Xi))\big)$

Proof. A normal form for (5.11) is easy to obtain and as follows:

$$(\forall^{\text{st}}k^0)(\exists^{\text{st}}N^0)(\forall \varepsilon, \varepsilon' \neq 0)\big(|\varepsilon|, |\varepsilon'| < \tfrac{1}{N} \to \Big|\tfrac{f(a+\varepsilon)-f(a)}{\varepsilon} - \tfrac{f(a+\varepsilon')-f(a)}{\varepsilon'}\Big| < \tfrac{1}{k}\big),$$

The proof is straightforward and analogous to the proof of Corollary 4.36. $\square$

Note that it is easy to define the derivate from the previous normal form of nonstandard differentiability by using $\mathsf{HAC}_{\mathsf{int}}$ to obtain a 'modulus of differentiability'.

5.4. **Classical computability theory.** In the previous sections, we established that theorems in Nonstandard Analysis give rise to results in computability theory like (4.36) or Theorem 4.11. Now, a distinction[27] exists between *higher-order* and *classical* computability theory. The latter (resp. the former) deals with computability on objects of type zero and one (resp. of any type) and it is clear that our above results are part of *higher-order* computability theory. In this section, we provide a generic example of how Nonstandard gives rise to results in *classical* computability theory. More examples (e.g. Ramsey's theorem and MCT) are treated in [83].

In this section, we study *König's lemma* (KOE for short), which is the statement that every finitely branching infinite tree has a path; KOE is equivalent to ACA_0 over RCA_0 by [87, III.7]. Let $\mathsf{KOE}_{\mathsf{ns}}$ be the statement that every *standard* tree as in KOE has a *standard* path. As expected, we shall prove Π^0_1-$\mathsf{TRANS} \leftrightarrow \mathsf{KOE}_{\mathsf{ns}}$ over P (Theorem 5.9), and obtain results in classical and higher-order computability theory from this equivalence (Theorems 5.10 and 5.11).

Theorem 5.9. *The system* P *proves* Π^0_1-$\mathsf{TRANS} \leftrightarrow \mathsf{KOE}_{\mathsf{ns}}$.

Proof. As noted before, we actually have an equivalence in (4.34), i.e. the standard functional μ^2 as in the latter allows us to decide existential formulas as long as a standard function describes the quantifier-free part as in the antecedent of (4.34). Now, if a standard tree T^1 is infinite and finitely branching, there is n^0 such that $(\forall m^0)(\exists \beta^{0^*})(\langle n\rangle * \beta \in T \wedge |\beta| = m)$. Since T is standard, the latter universal formula can be decided using μ^2 from (4.34). Say we have

$$(\forall^{\text{st}}n)\big[g(n) = 0 \leftrightarrow (\forall m^0)(\exists \beta^{0^*})(\langle n\rangle * \beta \in T \wedge |\beta| = m)\big]$$

[27]The distinction between 'higher-order' and 'classical' computability theory is not completely strict: Continuous functions on the real numbers are represented in second-order arithmetic by *type one* associates (aka Reverse Mathematics codes), as discussed in [56, §4].

where g is standard (and involves μ). Then $(\exists n^0)(g(n) = 0)$ implies $(\exists n^0 \leq \mu(g))(g(n) = 0)$, i.e. a search bounded by $\mu(g)$ provides a standard number n^0 such that the subtree of T starting with $\langle n\rangle$ is infinite. Since this subtree is also infinite, we can repeat this process to find a *standard* sequence α^1 such that $(\forall^{\rm st} n^0)(\overline{\alpha}n \in T)$. Applying Π^0_1-TRANS to the latter now yields $\mathsf{KOE}_{\mathsf{ns}}$.

Now assume $\mathsf{KOE}_{\mathsf{ns}}$ and suppose Π^0_1-TRANS is false, i.e. there is standard h^1 such that $(\forall^{\rm st} n^0)(h(n) = 0) \wedge (\exists m)(h(m) \neq 0)$. Define the tree T_0 as: $\sigma \in T_0$ if

$$(\forall i < |\sigma| - 1)(\sigma(i) = \sigma(i + 1)) \wedge h(\sigma(0)) \neq 0 \wedge (\forall i < \sigma(0))(h(i) = 0). \tag{5.12}$$

Clearly T_0 is standard, finitely branching, and infinite, but has no standard path, a contradiction. Hence Π^0_1-TRANS follows and we are done. $\square$

We refer to the previous proof as the 'textbook proof' of $\mathsf{KOE}_{\mathsf{ns}} \to \Pi^0_1$-$\mathsf{TRANS}$. The proof of this implication is indeed similar to the proof of $\mathsf{KOE} \to \mathsf{ACA}_0$ in Simpson's textbook on RM, as found in [87, III.7]. This 'textbook proof' is special in a specific technical sense, as will become clear below.

We first prove the associated higher-order result. Let $\mathsf{KOE}_{\mathsf{ef}}(t)$ be the statement that $t(T)$ is a path though T if the latter is an infinite and finitely branching tree.

Theorem 5.10. *From **any** proof of* $\mathsf{KOE}_{\mathsf{ns}} \leftrightarrow \Pi^0_1$-$\mathsf{TRANS}$ *in* P*, two terms s, u can be extracted such that* $\mathsf{E\text{-}PA}^\omega$ *proves:*

$$(\forall \mu^2)\big[\mathsf{MU}(\mu) \to \mathsf{KOE}_{\mathsf{ef}}(s(\mu))\big] \wedge (\forall t^{1\to 1})\big[\mathsf{KOE}_{\mathsf{ef}}(t) \to \mathsf{MU}(u(t))\big]. \tag{5.13}$$

Proof. We establish the second implication in (5.13) and leave the remaining one to the reader. To this end, note that $\mathsf{KOE}_{\mathsf{ns}} \to \Pi^0_1$-$\mathsf{TRANS}$ implies that

$$(\forall^{\rm st} T^1)(\exists^{\rm st} \alpha^1)A(T, \alpha) \to (\forall^{\rm st} f^1)(\exists^{\rm st} n)B(f, n), \tag{5.14}$$

where B is the formula in square brackets in (4.32) and $A(T, \alpha)$ is the *internal* formula expressing that α is a path in the finitely branching and infinite tree T. A standard functional provides standard output on standard input, and (5.14) yields

$$(\forall^{\rm st} t^{1\to 1})\big[(\forall T^1)A(T, t(T)) \to (\forall^{\rm st} f^1)(\exists^{\rm st} n)B(f, n)\big]. \tag{5.15}$$

Bringing outside all standard quantifiers, we obtain:

$$(\forall^{\rm st} t^{1\to 1}, f^1)(\exists^{\rm st} n^0)\big[(\forall T^1)A(T, t(T)) \to B(f, n)\big]. \tag{5.16}$$

Applying Theorem 4.2 to '$\mathsf{P} \vdash (5.16)$', we obtain a term u such that $\mathsf{E\text{-}PA}^\omega$ proves

$$(\forall t^{1\to 1}, f^1)(\exists n^0 \in u(t, f))\big[(\forall T^1)A(T, t(T)) \to B(f, n)\big]. \tag{5.17}$$

Taking the maximum of u, we obtain the second conjunct of (5.13). $\square$

To obtain the counterpart of the previous theorem in classical computability theory, consider the following 'second-order' version of (μ^2):

$$(\exists m, s)(\varphi^A_{e,s}(n) = m) \to (\exists m, s \leq \nu(e, n))(\varphi^A_{e,s}(n) = m). \tag{$\mathsf{MU}^A(\nu, e, n)$}$$

Furthermore, let $\mathsf{KOE}^A_{\mathsf{ef}}(t, e)$ be the statement that if an A-computable tree with index e^0 is infinite and finitely branching, then t^1 is a path through this tree.

Theorem 5.11. *From the 'textbook proof' of* $\mathsf{KOE}_{\mathsf{ns}} \to \Pi^0_1$-$\mathsf{TRANS}$ *in* P*, terms s, u can be extracted such that* $\mathsf{E\text{-}PA}^\omega$ *proves:*

$$(\forall e^0, n^0, C^1, \beta^1)\big[\mathsf{KOE}^C_{\mathsf{ef}}(\beta, u(e, n)) \to \mathsf{MU}^C(t(\beta, C, e, n), e, n)\big]. \tag{5.18}$$

Proof. We make essential use of the proofs of Theorems 5.9 and 5.10, and the associated notations. In particular, define the term $t^{1\to 1}$ by letting $t(h)$ be the tree T_0 from (5.12). Clearly, what is proved in the proof of Theorem 5.10 is in fact

$$(\forall^{\rm st} f^1)\big[(\exists^{\rm st}\alpha^1)A(t(f),\alpha)\to(\exists^{\rm st}n)B(f,n)\big], \tag{5.19}$$

where we used the notations from the proof of Theorem 5.10. Indeed, $\mathsf{KOE}_{\mathsf{ns}}$ is only applied for $T=t(f)=T_0$ in the proof of Theorem 5.10. Now (5.19) yields:

$$(\forall^{\rm st} f^1,\alpha^1)(\exists^{\rm st}n^0)\big[A(t(f),\alpha)\to B(f,n)\big]. \tag{5.20}$$

Applying Theorem 4.2 to '$\mathsf{P}\vdash(5.20)$', we obtain a term v such that $\mathsf{E\text{-}PA}^\omega$ proves

$$(\forall f^1,\alpha^1)(\exists n\in v(f,\alpha))\big[A(t(f),\alpha)\to B(f,n))\big],$$

and define $u(f,\alpha)$ as the maximum of all $v(f,\alpha)(i)$ for $i<|v(f,\alpha)|$. We obtain:

$$(\forall f^1,\alpha^1)\big[A(t(f),\alpha)\to B(f,u(f,\alpha)))\big], \tag{5.21}$$

We now modify (5.21) to obtain (5.18). To this end, define f_0^2 as follows: $f_0(e,n,C,k)=0$ if $(\exists m,s\le k)(\varphi^C_{e,s}(n)=m)$, and 1 otherwise. For this choice of function, namely taking $f^1=_1\lambda k.f_0$, the sentence (5.21) implies that

$$(\forall C^1,e^0,n^0,\alpha^1)\big[A(t(\lambda k.f_0),\alpha)\to B(\lambda k.f_0,u(\lambda k.f_0,\alpha)))\big]. \tag{5.22}$$

Now there are obvious (primitive recursive) terms x^1,v^1 such that for any finite sequence σ, we have $\sigma\in t(\lambda k.f_0)$ if and only if $\varphi^C_{x(e,n),v(e,n)}(\sigma)=1$; the definition of x^1,v^1 is implicit in the definition of t and f_0. Hence, with these terms, the antecedent and consequent of (5.22) are as required to yield (5.18). $\square$

Note that all objects in (5.18) are type zero or one, except the extracted terms. For those familiar with the concept, applying the so-called ECF-translation (See [97]) to (5.18) only changes these terms (to computable associates) and results in a 'pure' statement of second-order arithmetic, i.e. classical computability theory.

Let $\mathsf{KOE}_{\mathsf{ns}}(T)$ and $\Pi^0_1\text{-}\mathsf{TRANS}(f)$ be $\mathsf{KOE}_{\mathsf{ns}}$ and $\Pi^0_1\text{-}\mathsf{TRANS}$ without the quantifiers over f and T. The crux of the previous proof is that the 'textbook' proof of Theorem 5.9 establishes $(\forall^{\rm st}f^1)[\mathsf{KOE}_{\mathsf{ns}}(t(f))\to\Pi^0_1\text{-}\mathsf{TRANS}(f)]$ as in (5.19). In particular, the corresponding normal form (5.20) only involves *type zero and one* objects, rather than type two as in e.g. (4.38) for $\Pi^0_1\text{-}\mathsf{TRANS}\to\mathsf{MCT}_{\mathsf{ns}}$. Thus, *thanks to the textbook proof* all objects in (5.19) are 'of low enough type' to yield classical computability theory as in (5.18).

We are now ready to reveal the intended 'deeper' meaning of the term 'textbook proof': Intuitively, the latter refers to a proof (which may not exist) of an implication $(\forall^{\rm st}f)A(f)\to(\forall^{\rm st}g)B(g)$ which also establishes $(\forall^{\rm st}g)[A(t(g))\to B(g)]$, *and* in which the formula in square brackets has a normal form involving only standard quantifiers of type zero and one. Following the proof of Theorem 5.11, such a 'textbook proof' gives rise to results in **classical** computability theory. Such 'textbook proofs' also exist for MCT and Ramsey's theorem, as explored in [83].

In a nutshell, to obtain the previous theorem, one first establishes the 'nonstandard uniform' version (5.19) of $\mathsf{KOE}_{\mathsf{ns}}\to\Pi^0_1\text{-}\mathsf{TRANS}$, which yields the 'superpointwise' version (5.21). The latter is then modified to (5.22); this modification should be almost identical for other similar implications. In particular, versions of Theorems 5.11 and (5.19) are obtained in [83] for *König's lemma* and *Ramsey's* *theorem* ([87, III.7]). Similar or related results for the Reverse Mathematics of WKL_0 and ACA_0 should be straightforward to obtain.

In conclusion, higher-order computability results can be obtained from arbitrary proofs of $\mathsf{KOE}_{\mathsf{ns}} \to \Pi^0_1\text{-}\mathsf{TRANS}$, while the *textbook proof* as in the proof of Theorem 5.9 yields classical computability theory as in (5.18).

6. Nonstandard Analysis and intuitionistic mathematics

We have observed that the system P inhabits the twilight zone between the constructive and non-constructive: P is not the former as it (explicitly) includes the law of excluded middle, but P also has 'too many' constructive properties to be dismissed as merely the latter. In this section, we provide a possible explanation for the behaviour of P based on the foundational views of Brouwer and Troelstra.

6.1. **Nonstandard Analysis and Brouwer's view on logic.** In this section, we offer an alternative interpretation of our results based on Brouwer's claim that *logic is dependent on mathematics*. Indeed, Brouwer already explicitly stated in his dissertation that *logic depends upon mathematics* as follows:

> While thus mathematics is independent of logic, logic does depend upon mathematics: in the first place *intuitive logical reasoning* is that special kind of mathematical reasoning which remains if, considering mathematical structures, one restricts oneself to relations of whole and part; ([20, p. 127]; emphasis in the Dutch original)

According to Brouwer, logic is thus abstracted from mathematics, and he even found it conceivable that under different circumstances, a different abstraction would emerge from the same mathematics.

> Therefore it is easily conceivable that, given the same organization of the human intellect and consequently the same mathematics, a different language would have been formed, into which the language of logical reasoning, well known to us, would not fit. ([20, p. 129])

Building on Brouwer's viewpoint, it seems reasonable that different kinds of logical reasoning can emerge when the 'source', i.e. the underlying mathematics, is different. Indeed, Brouwer found it conceivable that a different logical language could emerge from the *same* mathematics, suggesting that changing mathematics in a fundamental way, a different kind of logical reasoning is (more?) likely to emerge.

This leads us to our alternative interpretation: If one fundamentally changes mathematics, as one arguably does when introducing Nonstandard Analysis, the logic will change along as the latter depends on the former in Brouwer's view. Since classical logic cannot really become 'more non-constructive', it stands to reason that the logic of (classical) Nonstandard Analysis actually is more constructive than plain classical logic. We present two strands of evidence for this claim.

First of all, we have observed in Section 4.1 that when introducing the notion of 'being standard' fundamental to Nonstandard Analysis, the law of excluded middle of classical mathematics moves from 'the original sin of non-constructivity' to a *computationally inert* principle which does not have any real non-constructive consequences anymore. In particular, adding LEM to $\mathsf{E\text{-}HA}^\omega$ leads to the classical system $\mathsf{E\text{-}PA}^\omega$, while adding LEM (and even $\mathsf{LEM}_{\mathsf{ns}}$) to H results in the system P *which has exactly the same term extraction theorem* as H by Theorems 4.1 and 4.2.

The latter even explicitly state that LEM does not influence the extracted term! In this way, our results on Nonstandard Analysis vindicate Brouwer's thesis that logic depends upon mathematics: Changing classical mathematics fundamentally by introducing the 'standard versus nonstandard' distinction from Nonstandard Analysis, the associated logic moves from classical logic to a new, more constructive, logic in which 'there exists a standard' has computational meaning akin to constructive mathematics and its BHK-interpretation.

As it happens, the observation that the introduction of the framework of Nonstandard Analysis changes the associated logic has been made by the pioneers of Nonstandard Analysis Robinson and Nelson. The latter has indeed qualified Nonstandard Analysis as a full-fledged new logic as follows.

> What Abraham Robinson invented is nothing less than a new logic. ([63, §1.6])

The previous claim is based on the following similar one by Robinson from his original monograph *Non-standard Analysis* ([74]).

> Returning now to the theory of this book, we observe that it is presented, naturally, within the framework of contemporary Mathematics, and thus appears to affirm the existence of all sorts of infinitary entities. However, from a formalist point of view we may look at our theory syntactically and may consider that what we have done is to introduce *new deductive procedures* rather than new mathematical entities. ([74, §10.7]; emphasis in original)

Thus, Robinson and Nelson already had the view that introducing the 'standard versus nonstandard' distinction from Nonstandard Analysis, the associated logic moves from classical logic to a new 'nonstandard' logic. Our goal in this paper has been to show that this new logic has constructive content.

Finally, a word of warning is in order: it is fair to say that both Robinson and Nelson had 'nonstandard' foundational views: Robinson subscribed to *formalism* ([75]) while Nelson even rejected the totality of the exponential function ([64]).

6.2. **Nonstandard Analysis and Troelstra's view of mathematics.** In this section, we offer an alternative interpretation of our results based on Troelstra's views on intuitionism.

First of all, Robinson-style Nonstandard Analysis involves a nonstandard model of a structure in which the latter is embedded, the textbook example being the real numbers $\mathbb{R}$ which form a subset of the 'hyperreal' numbers ${}^*\mathbb{R}$. At the risk of pointing out the obvious, one thus studies the real numbers $\mathbb{R}$ *from the outside*, namely as embedded into the (much) larger structure ${}^*\mathbb{R}$.

Secondly, Nelson-style Nonstandard Analysis as in IST does not introduce a new structure like ${}^*\mathbb{R}$, but imposes a new predicate 'is standard' onto the *existing* structure $\mathbb{R}$. As discussed above, the set of all standard real numbers cannot be formed in IST. Thus, IST is also fundamentally based on the 'standard versus nonstandard' distinction, but one does not view mathematics 'from the outside', but performs a kind of 'introspection' via the new standardness predicate.

Thirdly, while the above observations are not exactly new, the reader may be surprised to learn that Troelstra made similar claims regarding the nature of intuitionism, as follows.

> We may start with very simple concrete constructions, such as the natural numbers, and then gradually build up more complicated, but nevertheless "concrete" or "visualizable" structures. Finitism is concerned with such constructions only ([58, 3.4]). In intuitionism, we also want to exploit the idea that there is an intuitive concept of "constructive", by reflection on the properties of "constructions which are implicitely involved in the concept". (I.e. we attempt to discover principles by introspection.)
>
> Finitist constructions build up "from below"; reflecting on the general notion represents, so to speak, an approach "from the outside", "from above". ([96, §1.1])

In a nutshell, according to Troelstra, intuitionism deals with the study of concrete structures via a kind of 'introspection' or 'study from the outside'. Thus, we observe a remarkable similarity between the *methods of study* in intuitionism and Nonstandard Analysis: Nonstandard Analysis seems to originate from classical mathematics when applying to the latter the methodological techniques of intuitionism, namely 'introspection' (Nelson's IST) and 'study from the outside' (Robinsonian nonstandard models). Now, if all this sounds outrageous to the reader, we are not the first to make this claim: Wallet writes the following about Harthong-Reeb in [104, §7]:

> On the other hand, Harthong and Reeb explains in [40] that, far from being an artefact, nonstandard analysis necessarily results of an intuitionnistic [sic] interpretation of the classical mathematical formalism (see also [77]).

The references mentioned by Wallet are in French (and hard to come by), and we do not wish to go into details. Nonetheless, the message is loud and clear: There is a clear methodological connection between Nonstandard Analysis and intuitionism, as observed above and independently by Harthong-Reeb. Thus, as the French are wont to say: *Discutez!*

6.3. **Nonstandard Analysis and the awareness of time.** In our final section, we formulate a motivation for Nonstandard Analysis based on a fundamental intuitionist notion, namely the awareness of time. As to the latter's central status, Brouwer's *first act of intuitionism* reads as follows:

> Completely separating mathematics from mathematical language and hence from the phenomena of language described by theoretical logic, recognizing that intuitionistic mathematics is an essentially languageless activity of the mind having its origin in the perception of a move of time. This perception of a move of time may be described as the falling apart of a life moment into two distinct things, one of which gives way to the other, but is retained by memory. If the twoity thus born is divested of all quality, it passes into the empty form of the common substratum of all twoities. And it is this common substratum, this empty form, which is the basic intuition of mathematics. ([22], p. 4-5)

Needless to say, intuitionism is fundamentally rooted in the awareness of time. As such, the first act of intuitionism gives rise to the natural numbers (considered as a kind of potential infinity), while the second act (See [22, p. 8]) gives rise to the continuum. We shall not further discuss Brouwer's claims here, but only make the following observation regarding Nonstandard Analysis inspired by the first act. Like Brouwer, we assume the natural numbers to be given (by the first act or otherwise).

Observation 6.1 (Nonstandard Analysis and the awareness of time)**.** In intuitionism, the natural numbers (as a potential infinity) are obtained from Brouwer's fundamental concept *twoity* which is rooted in the awareness of time, in particular in the passage of time from one moment to the next. As a thought experiment, the reader should attempt to imagine a *long* period of time, say a day, a week, or even a decade, *as built up from moments of time* (say seconds).

While one *knowns* there are only finitely[28] many units of time which make up this long period of time, one cannot really *conceive* of the latter (as a whole) as built up from said units. In other words, despite the variable perception of time ('time flies' versus 'it took an eternity'), there are periods of time which cannot be grasped 'all at once' in terms of much smaller periods. Alternatively, it seems intuitively clear that consciously *experiencing* a unit of time cannot be done in less time; by extension *experiencing* a long period cannot be done in an instant or short time. Indeed, words like 'decade' are just that: words; they do not magically compress ten years of experience into one word, utterable within a second.

Thus, one observes an intuitive distinction between *short* and *long* time periods, i.e. ones which can be directly conceived as built up from units of time, and one which cannot. If a time period is short (resp. long) in this way, adjoining (resp. removing) another unit of time yields another short (resp. long) time period. If the distinction 'long versus short' is divested of all quality, one arrives at the usual notion of 'standard versus nonstandard' natural number from Nonstandard Analysis. Thus, the standardness predicate becomes grounded in one's awareness of time, namely in the intuitive distinction between long versus short periods of time.

By the previous observation, the fundamental notion 'standard versus nonstandard' from Nonstandard Analysis can also be grounded elegantly in one's awareness of time, *twoity* in particular. We leave it to the reader's own imagination what other properties (like e.g. Definition A.7) one can bestow upon the standardness predicate in the same way, i.e. motivated by the awareness of time. We point out that we avoided (the trap of) *ultrafinitism* (See [98, p. 29]) by considering the natural numbers as given.

In our opinion, Observation 6.1 is merely the 'next logical step' following Brouwer's first act inspired by Nonstandard Analysis. As such, this observation could have been made long time ago. Reasons for this apparent gap in the literature may be found in the rejection of Nonstandard Analysis by e.g. the constructivist Bishop, as discussed above.

Acknowledgement 6.2. This research was supported by the following entities: FWO Flanders, the John Templeton Foundation, the Alexander von Humboldt Foundation, LMU Munich (via the Excellence Initiative), and the Japan Society for the Promotion of Science. The work was done partially while the author was

[28]Apparently, there are 315,360,000 seconds in a decade.

visiting the Institute for Mathematical Sciences, National University of Singapore in 2016. The visit was supported by the Institute. The author expresses his gratitude towards these institutions. The author would like to thank Horst Osswald, Ulrich Kohlenbach, and Anil Nerode for their valuable advice.

Appendix A. The formal systems P and H

We introduce the systems P and H from [9] in detail in Section A.2 and A.3. To this end, we first introduce Gödel's system T in Section A.1.

A.1. **Gödel's system T.** In this section, we briefly introduce Gödel's system T and the associated systems $\mathsf{E\text{-}PA}^\omega$ and $\mathsf{E\text{-}PA}^{\omega *}$. In his famous *Dialectica* paper ([37]), Gödel defines an interpretation of intuitionistic arithmetic into a quantifier-free calculus of functionals. This calculus is now known as 'Gödel's system T', and is essentially just primitive recursive arithmetic ([23, §1.2.10]) with the schema of recursion expanded to *all finite types*. The set of all finite types $\mathbf{T}$ is:

$$\text{(i) } 0 \in \mathbf{T} \text{ and (ii) If } \sigma, \tau \in \mathbf{T} \text{ then } (\sigma \to \tau) \in \mathbf{T},$$

where 0 is the type of natural numbers, and $\sigma \to \tau$ is the type of mappings from objects of type σ to objects of type τ. Hence, Gödel's system T includes 'recursor' constants $\mathbf{R}^\rho$ for every finite type $\rho \in \mathbf{T}$, defining primitive recursion as follows:

$$\mathbf{R}^\rho(f, g, 0) := f \text{ and } \mathbf{R}^\rho(f, g, n+1) := g(n, \mathbf{R}^\rho(f, g, n)), \tag{PR}$$

for f^ρ and $g^{0\to(\rho\to\rho)}$. The system $\mathsf{E\text{-}PA}^\omega$ is a combination of *Peano Arithmetic* and system T, and the full axiom of extensionality (E). The detailed definition of $\mathsf{E\text{-}PA}^\omega$ may be found in [54, §3.3]; We do introduce the notion of equality and extensionality in $\mathsf{E\text{-}PA}^\omega$, as these notions are needed below.

Definition A.1. [Equality] The system $\mathsf{E\text{-}PA}^\omega$ includes equality between natural numbers '$=_0$' as a primitive. Equality '$=_\tau$' for type τ-objects x, y is then:

$$[x =_\tau y] \equiv (\forall z_1^{\tau_1} \dots z_k^{\tau_k})[x z_1 \dots z_k =_0 y z_1 \dots z_k] \tag{A.1}$$

if the type τ is composed as $\tau \equiv (\tau_1 \to \dots \to \tau_k \to 0)$. The usual inequality predicate '$\leq_0$' between numbers has an obvious definition, and the predicate '$\leq_\tau$' is just '$=_\tau$' with '$=_0$' replaced by '$\leq_0$' in (A.1). The *axiom of extensionality* is the statement that for all $\rho, \tau \in \mathbf{T}$, we have:

$$(\forall x^\rho, y^\rho, \varphi^{\rho\to\tau})\big[x =_\rho y \to \varphi(x) =_\tau \varphi(y)\big], \tag{E}$$

Next, we introduce $\mathsf{E\text{-}PA}^{\omega *}$, a definitional extension of $\mathsf{E\text{-}PA}^\omega$ from [9] with a type for finite sequences. In particular, the set $\mathbf{T}^*$ is defined as:

(i) $0 \in \mathbf{T}^*$, (ii) If $\sigma, \tau \in \mathbf{T}^*$ then $(\sigma \to \tau) \in \mathbf{T}^*$, and (iii) If $\sigma \in \mathbf{T}^*$, then $\sigma^* \in \mathbf{T}^*$,

where σ^* is the type of finite sequences of objects of type σ. The system $\mathsf{E\text{-}PA}^{\omega *}$ includes (PR) for all $\rho \in \mathbf{T}^*$, as well as dedicated 'list recursors' to handle finite sequences for any $\rho^* \in \mathbf{T}^*$. A detailed definition of $\mathsf{E\text{-}PA}^{\omega *}$ may be found in [9, §2.1]. We now introduce some notations specific to $\mathsf{E\text{-}PA}^{\omega *}$, as also used in [9].

Notation A.2 (Finite sequences)**.** The system $\mathsf{E\text{-}PA}^{\omega *}$ has a dedicated type for 'finite sequences of objects of type ρ', namely ρ^*. Since the usual coding of pairs of numbers goes through in $\mathsf{E\text{-}PA}^{\omega *}$, we shall not always distinguish between 0 and 0^*. Similarly, we do not always distinguish between 's^ρ' and '$\langle s^\rho \rangle$', where the former is 'the object s of type ρ', and the latter is 'the sequence of type ρ^* with only element s^ρ'. The empty sequence for the type ρ^* is denoted by '$\langle\rangle_\rho$', usually with the typing omitted. Furthermore, we denote by '$|s| = n$' the length of the finite sequence $s^{\rho^*} = \langle s_0^\rho, s_1^\rho, \dots, s_{n-1}^\rho \rangle$, where $|\langle\rangle| = 0$, i.e. the empty sequence has length zero. For sequences s^{ρ^*}, t^{ρ^*}, we denote by '$s * t$' the concatenation of s and t, i.e. $(s*t)(i) = s(i)$ for $i < |s|$ and $(s*t)(j) = t(|s|-j)$ for $|s| \le j < |s|+|t|$. For a sequence s^{ρ^*}, we define $\overline{s}N := \langle s(0), s(1), \dots, s(N)\rangle$ for $N^0 < |s|$. For a sequence $\alpha^{0\to\rho}$, we also write $\overline{\alpha}N = \langle \alpha(0), \alpha(1), \dots, \alpha(N)\rangle$ for *any* N^0. By way of shorthand, $q^\rho \in Q^{\rho^*}$ abbreviates $(\exists i < |Q|)(Q(i) =_\rho q)$. Finally, we shall use $\underline{x}, \underline{y}, \underline{t}, \dots$ as short for tuples $x_0^{\sigma_0}, \dots x_k^{\sigma_k}$ of possibly different type σ_i.

We have used $\mathsf{E\text{-}PA}^\omega$ and $\mathsf{E\text{-}PA}^{\omega*}$ interchangeably in this paper. Our motivation is the 'star morphism' used in Robinson's approach to Nonstandard Analysis, and the ensuing potential for confusion.

A.2. **The classical system P.** In this section, we introduce the system P, a conservative extension of $\mathsf{E\text{-}PA}^\omega$ with fragments of Nelson's IST.

To this end, we first introduce the base system $\mathsf{E\text{-}PA}^{\omega*}_{\mathrm{st}}$. We use the same definition as [9, Def. 6.1], where $\mathsf{E\text{-}PA}^{\omega*}$ is the definitional extension of $\mathsf{E\text{-}PA}^\omega$ with types for finite sequences as in [9, §2] and the previous section. The language of $\mathsf{E\text{-}PA}^{\omega*}_{\mathrm{st}}$ (and P) is the language of $\mathsf{E\text{-}PA}^{\omega*}$ extended with a new symbol 'st_ρ' for any finite type $\rho \in \mathbf{T}^*$ in the language of $\mathsf{E\text{-}PA}^{\omega*}$; The typing of 'st' is almost always omitted.

Definition A.3. The set $\mathcal{T}^*$ is defined as the collection of all the terms in the language of $\mathsf{E\text{-}PA}^{\omega*}$. The system $\mathsf{E\text{-}PA}^{\omega*}_{\mathrm{st}}$ is defined as $\mathsf{E\text{-}PA}^{\omega*} + \mathcal{T}^*_{\mathrm{st}} + \mathsf{IA}^{\mathrm{st}}$, where $\mathcal{T}^*_{\mathrm{st}}$ consists of the following axiom schemas.

(i) The schema[29] $\mathrm{st}(x) \wedge x = y \to \mathrm{st}(y)$,
(ii) The schema providing for each closed[30] term $t \in \mathsf{T}^*$ the axiom $\mathrm{st}(t)$.
(iii) The schema $\mathrm{st}(f) \wedge \mathrm{st}(x) \to \mathrm{st}(f(x))$.

The external induction axiom $\mathsf{IA}^{\mathrm{st}}$ is as follows.

$$\Phi(0) \wedge (\forall^{\mathrm{st}} n^0)(\Phi(n) \to \Phi(n+1)) \to (\forall^{\mathrm{st}} n^0)\Phi(n). \tag{$\mathsf{IA}^{\mathrm{st}}$}$$

Secondly, we introduce some essential fragments of IST studied in [9].

Definition A.4. [External axioms of P]

(1) $\mathsf{HAC}_{\mathsf{int}}$: For any internal formula φ, we have

$$(\forall^{\mathrm{st}} x^\rho)(\exists^{\mathrm{st}} y^\tau)\varphi(x,y) \to \big(\exists^{\mathrm{st}} F^{\rho\to\tau^*}\big)(\forall^{\mathrm{st}} x^\rho)(\exists y^\tau \in F(x))\varphi(x,y), \tag{A.2}$$

(2) I: For any internal formula φ, we have

$$(\forall^{\mathrm{st}} x^{\sigma^*})(\exists y^\tau)(\forall z^\sigma \in x)\varphi(z,y) \to (\exists y^\tau)(\forall^{\mathrm{st}} x^\sigma)\varphi(x,y),$$

(3) The system P is $\mathsf{E\text{-}PA}^{\omega*}_{\mathrm{st}} + \mathsf{I} + \mathsf{HAC}_{\mathsf{int}}$.

Note that I and $\mathsf{HAC}_{\mathsf{int}}$ are fragments of Nelson's axioms *Idealisation* and *Standard part*. By definition, F in (A.2) only provides a *finite sequence* of witnesses to $(\exists^{\mathrm{st}} y)$, explaining its name *Herbrandized Axiom of Choice*.

[29] The language of $\mathsf{E\text{-}PA}^{\omega*}_{\mathrm{st}}$ contains a symbol st_σ for each finite type σ, but the subscript is essentially always omitted. Hence $\mathsf{T}^*_{\mathrm{st}}$ is an *axiom schema* and not an axiom.

[30] A term is called *closed* in [9] (and in this paper) if all variables are bound via lambda abstraction. Thus, if $\underline{x}, \underline{y}$ are the only variables occurring in the term t, the term $(\lambda \underline{x})(\lambda \underline{y})t(\underline{x}, \underline{y})$ is closed while $(\lambda \underline{x})t(\underline{x}, \underline{y})$ is not. The second axiom in Definition A.3 thus expresses that $\mathrm{st}_\tau\big((\lambda \underline{x})(\lambda \underline{y})t(\underline{x}, \underline{y})\big)$ if $(\lambda \underline{x})(\lambda \underline{y})t(\underline{x}, \underline{y})$ is of type τ. We usually omit lambda abstraction for brevity.

The system P is connected to $\mathsf{E\text{-}PA}^\omega$ by the following theorem. Here, the superscript '$S_{\rm st}$' is the syntactic translation defined in [9, Def. 7.1].

Theorem A.5. *Let $\Phi(\underline{a})$ be a formula in the language of $\mathsf{E\text{-}PA}^{\omega*}_{\rm st}$ and suppose $\Phi(\underline{a})^{S_{\rm st}} \equiv \forall^{\rm st}\underline{x}\,\exists^{\rm st}\underline{y}\,\varphi(\underline{x},\underline{y},\underline{a})$. If Δ_{int} is a collection of internal formulas and*

$$\mathsf{P} + \Delta_{\mathsf{int}} \vdash \Phi(\underline{a}), \tag{A.3}$$

then one can extract from the proof a sequence of closed[31] *terms t in $\mathcal{T}^*$ such that*

$$\mathsf{E\text{-}PA}^{\omega*} + \Delta_{\mathsf{int}} \vdash\ \forall\underline{x}\,\exists\underline{y} \in \underline{t}(\underline{x})\ \varphi(\underline{x},\underline{y},\underline{a}). \tag{A.4}$$

Proof. Immediate by [9, Theorem 7.7]. □

The proofs of the soundness theorems in [9, §5-7] provide an algorithm to obtain the term t from the theorem. In particular, these terms can be 'read off' from the nonstandard proofs. The translation $S_{\rm st}$ can be formalised in any reasonable[32] system of constructive mathematics. In fact, the formalisation of the results in [9] in the proof assistant Agda (based on Martin-Löf's constructive type theory [60]) is underway in [109].

In light of the above results and those in [81], the following corollary (which is not present in [9]) is essential to our results. Indeed, the following corollary expresses that we may obtain effective results as in (A.6) from any theorem of Nonstandard Analysis which has the same form as in (A.5). It was shown in [79, 81, 85] that the scope of this corollary includes the Big Five systems of Reverse Mathematics and the associated 'zoo' ([29]).

Corollary A.6. *If Δ_{int} is a collection of internal formulas and ψ is internal, and*

$$\mathsf{P} + \Delta_{\mathsf{int}} \vdash (\forall^{\rm st}\underline{x})(\exists^{\rm st}\underline{y})\psi(\underline{x},\underline{y},\underline{a}), \tag{A.5}$$

then one can extract from the proof a sequence of closed[33] *terms t in $\mathcal{T}^*$ such that*

$$\mathsf{E\text{-}PA}^{\omega*} + \Delta_{\mathsf{int}} \vdash (\forall\underline{x})(\exists\underline{y} \in t(\underline{x}))\psi(\underline{x},\underline{y},\underline{a}). \tag{A.6}$$

Proof. Clearly, if for internal ψ and $\Phi(\underline{a}) \equiv (\forall^{\rm st}\underline{x})(\exists^{\rm st}\underline{y})\psi(x,y,a)$, we have $[\Phi(\underline{a})]^{S_{\rm st}} \equiv \Phi(\underline{a})$, then the corollary follows immediately from the theorem. A tedious but straightforward verification using the clauses (i)-(v) in [9, Def. 7.1] establishes that indeed $\Phi(\underline{a})^{S_{\rm st}} \equiv \Phi(\underline{a})$. We undertake this verification in Section A.5 below. □

A.3. **The constructive system H.** In this section, we define the system H, the constructive counterpart of P. The system H was first introduced in [9, §5.2], and constitutes a conservative extension of Heyting arithmetic $\mathsf{E\text{-}HA}^\omega$ by [9, Cor. 5.6]. We now study the system H in more detail.

Similar to Definition A.3, we define $\mathsf{E\text{-}HA}^{\omega*}_{\rm st}$ as $\mathsf{E\text{-}HA}^{\omega*} + \mathsf{T}^*_{\rm st} + \mathsf{IA}^{\rm st}$, where $\mathsf{E\text{-}HA}^{\omega*}$ is just $\mathsf{E\text{-}PA}^{\omega*}$ without the law of excluded middle. Furthermore, we define

$$\mathsf{H} \equiv \mathsf{E\text{-}HA}^{\omega*}_{\rm st} + \mathsf{HAC} + \mathsf{I} + \mathsf{NCR} + \mathsf{HIP}_{\forall^{\rm st}} + \mathsf{HGMP}^{\rm st},$$

where HAC is $\mathsf{HAC}_{\mathsf{int}}$ without any restriction on the formula, and where the remaining axioms are defined in the following definition.

[31] Recall the definition of closed terms from [9] as sketched in Footnote 30.

[32] Here, a 'reasonable' system is one which can prove the usual properties of finite lists, for which the presence of the exponential function suffices.

[33] Recall the definition of closed terms from [9] as sketched in Footnote 30.

Definition A.7. [Three axioms of H]

(1) $\mathsf{HIP}_{\forall^{\rm st}}$

$$[(\forall^{\rm st}x)\phi(x)\to(\exists^{\rm st}y)\Psi(y)]\to(\exists^{\rm st}y')[(\forall^{\rm st}x)\phi(x)\to(\exists y\in y')\Psi(y)],$$

where $\Psi(y)$ is any formula and $\phi(x)$ is an internal formula of $\mathsf{E\text{-}HA}^{\omega*}$.

(2) $\mathsf{HGMP}^{\rm st}$

$$[(\forall^{\rm st}x)\phi(x)\to\psi]\to(\exists^{\rm st}x')[(\forall x\in x')\phi(x)\to\psi]$$

where $\phi(x)$ and ψ are internal formulas in the language of $\mathsf{E\text{-}HA}^{\omega*}$.

(3) NCR

$$(\forall y^\tau)(\exists^{\rm st}x^\rho)\Phi(x,y)\to(\exists^{\rm st}x^{\rho^*})(\forall y^\tau)(\exists x'\in x)\Phi(x',y),$$

where Φ is any formula of $\mathsf{E\text{-}HA}^{\omega*}$

Intuitively speaking, the first two axioms of Definition A.7 allow us to perform a number of *non-constructive operations* (namely *Markov's principle* and *independence of premises*) on the standard objects of the system H, provided we introduce a 'Herbrandisation' as in the consequent of HAC, i.e. a finite list of possible witnesses rather than one single witness. Furthermore, while H includes idealisation I, one often uses the latter's *classical contraposition*, explaining why NCR is useful (and even essential) in the context of intuitionistic logic.

Surprisingly, the axioms from Definition A.7 are exactly what is needed to convert nonstandard definitions (of continuity, integrability, convergence, et cetera) into the normal form $(\forall^{\rm st}x)(\exists^{\rm st}y)\varphi(x,y)$ for internal φ, as is clear from Theorem 5.5. This normal form plays an equally important role in the constructive case as in the classical case by the following theorem.

Theorem A.8. *If Δ_{int} is a collection of internal formulas, φ is internal, and*

$$\mathsf{H}+\Delta_{\mathsf{int}}\vdash\forall^{\rm st}\underline{x}\,\exists^{\rm st}\underline{y}\,\varphi(\underline{x},\underline{y},\underline{a}),\tag{A.7}$$

then one can extract from the proof a sequence of closed[34] *terms t in $\mathcal{T}^*$ such that*

$$\mathsf{E\text{-}HA}^{\omega*}+\Delta_{\mathsf{int}}\vdash\ \forall\underline{x}\,\exists\underline{y}\in\underline{t}(\underline{x})\ \varphi(\underline{x},\underline{y},\underline{a}).\tag{A.8}$$

Proof. Immediate by [9, Theorem 5.9]. □

The proofs of the soundness theorems in [9, §5-7] provide an algorithm to obtain the term t from the theorem. Finally, we point out one very useful principle to which we have access.

Theorem A.9. *The systems* P,H *prove* overspill, *i.e.*

$$(\forall^{\rm st}x^\rho)\varphi(x)\to(\exists y^\rho)\big[\neg{\rm st}(y)\wedge\varphi(y)\big],\tag{OS}$$

for any internal formula φ.

Proof. See [9, Prop. 3.3]. □

In conclusion, we have introduced the systems H and P which are conservative extensions of Peano and Heyting arithmetic with fragments of Nelson's internal set theory. We have observed that central to the conservation results (Corollary A.6 and Theorem A.5) is the normal form $(\forall^{\rm st}x)(\exists^{\rm st}y)\varphi(x,y)$ for internal φ.

[34] Recall the definition of closed terms from [9] as sketched in Footnote 30.

A.4. **Notations and definitions in H and P.** In this section, we introduce notations relating and defintions to H and P.

First of all, we mostly use the same notations as in [9].

Remark A.10 (Notations)**.** We write $(\forall^{\text{st}} x^\tau)\Phi(x^\tau)$ and $(\exists^{\text{st}} x^\sigma)\Psi(x^\sigma)$ as short for $(\forall x^\tau)\big[\text{st}(x^\tau) \to \Phi(x^\tau)\big]$ and $(\exists x^\sigma)\big[\text{st}(x^\sigma) \wedge \Psi(x^\sigma)\big]$. A formula A is 'internal' if it does not involve st, and external otherwise. The formula A^{st} is defined from A by appending 'st' to all quantifiers (except bounded number quantifiers).

Secondly, we will use the usual notations for natural, rational and real numbers and functions as introduced in [55, p. 288-289]. (and [87, I.8.1] for the former). We only list the definition of real number and related notions in P and related systems.

Definition A.11 (Real numbers and related notions)**.**

(1) Natural numbers correspond to type zero objects, and we use 'n^0' and '$n \in \mathbb{N}$' interchangeably. Rational numbers are defined in the usual way as quotients of natural numbers, and '$q \in \mathbb{Q}$' has its usual meaning.
(2) A (standard) real number x is a (standard) fast-converging Cauchy sequence $q^1_{(\cdot)}$, i.e. $(\forall n^0, i^0)(|q_n - q_{n+i})| <_0 \frac{1}{2^n})$. We use Kohlenbach's 'hat function' from [55, p. 289] to guarantee that every sequence f^1 is a real.
(3) We write '$x \in \mathbb{R}$' to express that $x^1 = (q^1_{(\cdot)})$ is a real as in the previous item and $[x](k) := q_k$ for the k-th approximation of x.
(4) Two reals x, y represented by $q_{(\cdot)}$ and $r_{(\cdot)}$ are *equal*, denoted $x =_{\mathbb{R}} y$, if $(\forall n^0)(|q_n - r_n| \leq \frac{1}{2^n})$. Inequality $<_{\mathbb{R}}$ is defined similarly.
(5) We write $x \approx y$ if $(\forall^{\text{st}} n^0)(|q_n - r_n| \leq \frac{1}{2^n})$ and $x \gg y$ if $x > y \wedge x \not\approx y$.
(6) Functions $F : \mathbb{R} \to \mathbb{R}$ mapping reals to reals are represented by functionals $\Phi^{1\to 1}$ mapping equal reals to equal reals, i.e.
$$(\forall x, y)(x =_{\mathbb{R}} y \to \Phi(x) =_{\mathbb{R}} \Phi(y)). \tag{RE}$$
(7) Sets of objects of type ρ are denoted $X^{\rho\to 0}, Y^{\rho\to}, Z^{\rho\to 0}, \dots$ and are given by their characteristic functions $f_X^{\rho\to 0}$, i.e. $(\forall x^\rho)[x \in X \leftrightarrow f_X(x) =_0 1]$, where $f_X^{\rho\to 0}$ is assumed to output zero or one.

Thirdly, we use the usual extensional notion of equality.

Remark A.12 (Equality)**.** The systems P and H include equality between natural numbers '$=_0$' as a primitive. Equality '$=_\tau$' for type τ-objects x, y is defined as:
$$[x =_\tau y] \equiv (\forall z_1^{\tau_1} \dots z_k^{\tau_k})[x z_1 \dots z_k =_0 y z_1 \dots z_k] \tag{A.9}$$
if the type τ is composed as $\tau \equiv (\tau_1 \to \dots \to \tau_k \to 0)$. Inequality '$\leq_\tau$' is then just (A.1) with '$=_0$' replaced by '$\leq_0$'. In the spirit of Nonstandard Analysis, we define 'approximate equality $\approx_\tau$' as follows:
$$[x \approx_\tau y] \equiv (\forall^{\text{st}} z_1^{\tau_1} \dots z_k^{\tau_k})[x z_1 \dots z_k =_0 y z_1 \dots z_k] \tag{A.10}$$
with the type τ as above. All the above systems include the *axiom of extensionality*:
$$(\forall x^\rho, y^\rho, \varphi^{\rho\to\tau})\big[x =_\rho y \to \varphi(x) =_\tau \varphi(y)\big]. \tag{E}$$
However, as noted in [9, p. 1973], the so-called axiom of *standard* extensionality $(\mathsf{E})^{\text{st}}$ is problematic and cannot be included in P or H. In particular, we cannot have term extraction as in Theorem 4.2 in the presence of this external axiom.

Finally, we note that equality on the reals '$=_{\mathbb{R}}$' is a defined notion and not an equality '$=_\tau$' of P. Hence, the first item of Definition A.3 **does not** apply to '$=_{\mathbb{R}}$'. In fact, $\mathsf{P} + (\forall x, y \in \mathbb{R})((x =_{\mathbb{R}} y \wedge \text{st}(x)) \to \text{st}(y))$ proves[35] a contradiction.

A.5. **Technical results.** In this section, we have gathered the proofs of some of the above theorems too lengthy to fit the body of the text.

First of all, for completeness, we prove our 'term extraction' result in detail.

Theorem A.13. *If Δ_{int} is a collection of internal formulas and ψ is internal, and*

$$\mathsf{P} + \Delta_{\mathsf{int}} \vdash (\forall^{\text{st}} \underline{x})(\exists^{\text{st}} \underline{y})\psi(\underline{x}, \underline{y}, \underline{a}), \tag{A.11}$$

then one can extract from the proof a sequence of closed[36] *terms t in $\mathcal{T}^*$ such that*

$$\mathsf{E\text{-}PA}^{\omega*} + \mathsf{QF\text{-}AC}^{1,0} + \Delta_{\mathsf{int}} \vdash (\forall \underline{x})(\exists \underline{y} \in t(\underline{x}))\psi(\underline{x}, \underline{y}, \underline{a}). \tag{A.12}$$

Proof. Clearly, if for internal ψ and $\Phi(\underline{a}) \equiv (\forall^{\text{st}} \underline{x})(\exists^{\text{st}} \underline{y})\psi(x, y, a)$, we have $[\Phi(\underline{a})]^{S_{\text{st}}} \equiv \Phi(\underline{a})$, then the corollary follows immediately from Theorem A.5. A tedious but straightforward verification using the clauses (i)-(v) in [9, Def. 7.1] establishes that indeed $\Phi(\underline{a})^{S_{\text{st}}} \equiv \Phi(\underline{a})$. For completeness, we now list these five inductive clauses and perform this verification.

Hence, suppose $\Phi(\underline{a})$ and $\Psi(\underline{b})$ in the language of P have the interpretations

$$\Phi(\underline{a})^{S_{\text{st}}} \equiv (\forall^{\text{st}} \underline{x})(\exists^{\text{st}} \underline{y})\varphi(\underline{x}, \underline{y}, \underline{a}) \text{ and } \Psi(\underline{b})^{S_{\text{st}}} \equiv (\forall^{\text{st}} \underline{u})(\exists^{\text{st}} \underline{v})\psi(\underline{u}, \underline{v}, \underline{b}), \tag{A.13}$$

for internal ψ, φ. These formulas then behave as follows by [9, Def. 7.1]:

(i) $\psi^{S_{\text{st}}} := \psi$ for atomic internal ψ.
(ii) $\big(\text{st}(z)\big)^{S_{\text{st}}} := (\exists^{\text{st}} x)(z = x)$.
(iii) $(\neg\Phi)^{S_{\text{st}}} := (\forall^{\text{st}} \underline{Y})(\exists^{\text{st}} \underline{x})(\forall \underline{y} \in \underline{Y}[\underline{x}])\neg\varphi(\underline{x}, \underline{y}, \underline{a})$.
(iv) $(\Phi \vee \Psi)^{S_{\text{st}}} := (\forall^{\text{st}} \underline{x}, \underline{u})(\exists^{\text{st}} \underline{y}, \underline{v})[\varphi(\underline{x}, \underline{y}, \underline{a}) \vee \psi(\underline{u}, \underline{v}, \underline{b})]$
(v) $\big((\forall z)\Phi\big)^{S_{\text{st}}} := (\forall^{\text{st}} \underline{x})(\exists^{\text{st}} \underline{y})(\forall z)(\exists \underline{y}' \in \underline{y})\varphi(\underline{x}, \underline{y}', z)$

Hence, fix $\Phi_0(\underline{a}) \equiv (\forall^{\text{st}} \underline{x})(\exists^{\text{st}} \underline{y})\psi_0(\underline{x}, \underline{y}, \underline{a})$ with internal ψ_0, and note that $\phi^{S_{\text{st}}} \equiv \phi$ for any internal formula. We have $[\text{st}(\underline{y})]^{S_{\text{st}}} \equiv (\exists^{\text{st}} \underline{w})(\underline{w} = \underline{y})$ and also

$$[\neg\text{st}(\underline{y})]^{S_{\text{st}}} \equiv (\forall^{\text{st}} \underline{W})(\exists^{\text{st}} \underline{x})(\forall \underline{w} \in \underline{W}[\underline{x}])\neg(\underline{w} = \underline{y}) \equiv (\forall^{\text{st}} \underline{w})(\underline{w} \neq \underline{y}).$$

Hence, $[\neg\text{st}(\underline{y}) \vee \neg\psi_0(\underline{x}, \underline{y}, \underline{a})]^{S_{\text{st}}}$ is just $(\forall^{\text{st}} \underline{w})[(\underline{w} \neq \underline{y}) \vee \neg\psi_0(\underline{x}, \underline{y}, \underline{a})]$, and

$$\big[(\forall \underline{y})[\neg\text{st}(\underline{y}) \vee \neg\psi_0(\underline{x}, \underline{y}, \underline{a})]\big]^{S_{\text{st}}} \equiv (\forall^{\text{st}} \underline{w})(\exists^{\text{st}} \underline{v})(\forall \underline{y})(\exists \underline{v}' \in \underline{v})[\underline{w} \neq \underline{y} \vee \neg\psi_0(\underline{x}, \underline{y}, \underline{a})].$$

which is just $(\forall^{\text{st}} \underline{w})(\forall \underline{y})[(\underline{w} \neq \underline{y}) \vee \neg\psi_0(\underline{x}, \underline{y}, \underline{a})]$. Furthermore, we have

$$\begin{aligned}\big[(\exists^{\text{st}} y)\psi_0(\underline{x}, \underline{y}, \underline{a})\big]^{S_{\text{st}}} &\equiv \big[\neg(\forall \underline{y})[\neg\text{st}(\underline{y}) \vee \neg\psi_0(\underline{x}, \underline{y}, \underline{a})]\big]^{S_{\text{st}}} \\ &\equiv (\forall^{\text{st}} \underline{V})(\exists^{\text{st}} \underline{w})(\forall \underline{v} \in \underline{V}[\underline{w}])\neg[(\forall \underline{y})[(\underline{w} \neq \underline{y}) \vee \neg\psi_0(\underline{x}, \underline{y}, \underline{a})]]. \\ &\equiv (\exists^{\text{st}} \underline{w})(\exists \underline{y})[(\underline{w} = \underline{y}) \wedge \psi_0(\underline{x}, \underline{y}, \underline{a})]] \equiv (\exists^{\text{st}} \underline{w})\psi_0(\underline{x}, \underline{w}, \underline{a}).\end{aligned}$$

Hence, we have proved so far that $(\exists^{\text{st}} \underline{y})\psi_0(\underline{x}, \underline{y}, \underline{a})$ is invariant under S_{st}. By the previous, we also obtain:

$$\big[\neg\text{st}(\underline{x}) \vee (\exists^{\text{st}} y)\psi_0(\underline{x}, \underline{y}, \underline{a})\big]^{S_{\text{st}}} \equiv (\forall^{\text{st}} \underline{w}')(\exists^{\text{st}} \underline{w})[(\underline{w}' \neq \underline{x}) \vee \psi_0(\underline{x}, \underline{w}, \underline{a})].$$

---

[35] Fix nonstandard N^0 and define $x =_1 0$ and $y := (r_{(\cdot)})$ where $r_k = \frac{1}{2^{k+N}}$. Then $x =_{\mathbb{R}} y$ and $\text{st}_1(x)$ while r_0 clearly is nonstandard.

[36] Recall the definition of closed terms from [9] as sketched in Footnote 30.

Our final computation now yields the desired result:

$$\begin{aligned}\big[(\forall^{\rm st}\underline{x})(\exists^{\rm st}y)\psi_0(\underline{x},\underline{y},\underline{a})\big]^{S_{\rm st}} &\equiv \big[(\forall\underline{x})(\neg{\rm st}(\underline{x})\vee(\exists^{\rm st}y)\psi_0(\underline{x},\underline{y},\underline{a}))\big]^{S_{\rm st}}\\ &\equiv (\forall^{\rm st}\underline{w}')(\exists^{\rm st}\underline{w})(\forall\underline{x})(\exists\underline{w}''\in\underline{w})[(\underline{w}'\neq\underline{x})\vee\psi_0(\underline{x},\underline{w}'',\underline{a})].\\ &\equiv (\forall^{\rm st}\underline{w}')(\exists^{\rm st}\underline{w})(\exists\underline{w}''\in\underline{w})\psi_0(\underline{w}',\underline{w}'',\underline{a}).\end{aligned}$$

The last step is obtained by taking $\underline{x}=\underline{w}'$. Hence, we may conclude that the normal form $(\forall^{\rm st}\underline{x})(\exists^{\rm st}y)\psi_0(\underline{x},\underline{y},\underline{a})$ is invariant under $S_{\rm st}$, and we are done. □

Next, we prove two theorems which provide a normal form for STP and establishes the latter's relationship with the special and intuitionistic fan functional. The function g^1 from (A.15) is called a *standard part* of f^1.

Theorem A.14. *In* P*, both* STP *and* $\mathsf{STP}_{\mathbb{R}}$ *are equivalent to either of the following:*

$$\begin{aligned}(\forall^{\rm st}g^2)(\exists^{\rm st}w)\big[(\forall T^1\leq_1 1)\big((\forall\alpha^1\in w(1))(\overline{\alpha}g(\alpha)\not\in T)\\ \to(\forall\beta\leq_1 1)(\exists i^0\leq w(2))(\overline{\beta}i\not\in T)\big)\big],\end{aligned}\tag{A.14}$$

$$(\forall f^1)(\exists^{\rm st}g^1)\big((\forall^{\rm st}n^0)(\exists^{\rm st}m^0)(f(n)=m)\to f\approx_1 g\big).\tag{A.15}$$

Furthermore, P *proves* $(\exists^{\rm st}\Theta)\mathsf{SCF}(\Theta)\to\mathsf{STP}$.

Proof. First of all, since any individual real can be given a binary representation (See [43]), the equivalence between STP and $\mathsf{STP}_{\mathbb{R}}$ is immediate. Next, STP is easily seen to be equivalent to

$$\begin{aligned}(\forall T^1\leq_1 1)\big[(\forall^{\rm st}n)(\exists\beta^0)(|\beta|=n\wedge\beta\in T)\\ \to(\exists^{\rm st}\alpha^1\leq_1 1)(\forall^{\rm st}n^0)(\overline{\alpha}n\in T)\big],\end{aligned}\tag{A.16}$$

and this equivalence may also be found in [78, Theorem 3.2]. For completeness, we first prove the equivalence $\mathsf{STP}\leftrightarrow$ (A.16). Assume STP and apply overspill to $(\forall^{\rm st}n)(\exists\beta^0)(|\beta|=n\wedge\beta\in T)$ to obtain $\beta_0^0\in T$ with nonstandard length $|\beta_0|$. Now apply STP to $\beta^1:=\beta_0*00\dots$ to obtain a *standard* $\alpha^1\leq_1 1$ such that $\alpha\approx_1\beta$ and hence $(\forall^{\rm st}n)(\overline{\alpha}n\in T)$. For the reverse direction, let f^1 be a binary sequence, and define a binary tree T_f which contains all initial segments of f. Now apply (A.16) for $T=T_f$ to obtain STP.

For the implication (A.14)→(A.16), note that (A.14) implies for all standard g^2

$$(\forall T^1\leq_1 1)(\exists^{\rm st}(\alpha^1\leq_1 1,\,k^0)\big[(\overline{\alpha}g(\alpha)\not\in T)\to(\forall\beta\leq_1 1)(\exists i\leq k)(\overline{\beta}i\not\in T)\big],\tag{A.17}$$

which in turn yields, by bringing all standard quantifiers inside again, that:

$$(\forall T\leq_1 1)\big[(\exists^{\rm st}g^2)(\forall^{\rm st}\alpha\leq_1 1)(\overline{\alpha}g(\alpha)\not\in T)\to(\exists^{\rm st}k)(\forall\beta\leq_1 1)(\overline{\beta}k\not\in T)\big],\tag{A.18}$$

To obtain (A.16) from (A.18), apply $\mathsf{HAC_{int}}$ to $(\forall^{\rm st}\alpha^1\leq_1 1)(\exists^{\rm st}n)(\overline{\alpha}n\not\in T)$ to obtain standard $\Psi^{1\to 0^*}$ with $(\forall^{\rm st}\alpha^1\leq_1 1)(\exists n\in\Psi(\alpha))(\overline{\alpha}n\not\in T)$, and defining $g(\alpha):=\max_{i<|\Psi|}\Psi(\alpha)(i)$ we obtain g as in the antecedent of (A.18). We obtain:

$$(\forall T^1\leq_1 1)\big[(\forall^{\rm st}\alpha^1\leq_1 1)(\exists^{\rm st}n)(\overline{\alpha}n\not\in T)\to(\exists^{\rm st}k)(\forall\beta\leq_1 1)(\overline{\beta}i\not\in T)\big],\tag{A.19}$$

which is the contraposition of (A.16), using classical logic. For the implication (A.16) → (A.14), consider the contraposition of (A.16), i.e. (A.19), and note that the latter implies (A.18). Now push all standard quantifiers outside as follows:

$$(\forall^{\rm st}g^2)(\forall T^1\leq_1 1)(\exists^{\rm st}(\alpha^1\leq_1 1,\ k^0)\big[(\overline{\alpha}g(\alpha)\not\in T)\to(\forall\beta\leq_1 1)(\exists i\leq k)(\overline{\beta}i\not\in T)\big],$$

and applying idealisation I yields (A.14) by taking the maximum of all elements pertaining to k. The equivalence involving (A.14) also immediately establishes the final part of the theorem.

For the remaining equivalence, the implication (A.15) $\to$ STP is trivial, and for the reverse implication, fix f^1 such that $(\forall^{\text{st}} n)(\exists^{\text{st}} m) f(n) = m$ and let h^1 be such that $(\forall n, m)(f(n) = m \leftrightarrow h(n,m) = 1)$. Applying $\mathsf{HAC}_{\mathsf{int}}$ to the former formula, there is standard $\Phi^{0\to 0^*}$ such that $(\forall^{\text{st}} n)(\exists m \in \Phi(n)) f(n) = m$, and define $\Psi(n) := \max_{i<|\Phi(n)|} \Phi(n)(i)$. Now define the sequence $\alpha_0 \leq_1 1$ as follows: $\alpha_0(0) := h(0,0)$, $\alpha_0(1) := h(0,1)$, $\ldots$, $\alpha_0(\Psi(0)) := h(0,\Psi(0))$, $\alpha_0(\Psi(0)+1) := h(1,0)$, $\alpha_0(\Psi(0)+2) := h(1,1)$, $\ldots$, $\alpha_0(\Psi(0)+\Psi(1)) := h(1,\Psi(1))$, et cetera. Now let $\beta_0^1 \leq_1 1$ be the standard part of α_0 provided by STP and define $g(n) := (\mu m \leq \Psi(n))\big[\beta_0(\sum_{i=0}^{n-1} \Psi(i) + m) = 1\big]$. By definition, g^1 is standard and $f \approx_1 g$. $\square$

Theorem A.15. *The axiom* STP *can be proved in* P *plus the axiom*

$$(\forall^{\text{st}} Y^2)(\forall f^1, g^1 \leq_1 1)(f \approx_1 g \to Y(f) =_0 Y(g)). \tag{NUC}$$

Proof. Resolving '$\approx_1$', NUC implies that

$$(\forall^{\text{st}} Y^2)(\forall f^1, g^1 \leq_1 1)(\exists^{\text{st}} N^0)(\overline{f}N =_0 \overline{g}N \to Y(f) =_0 Y(g)). \tag{A.20}$$

Applying *Idealisation* I to (A.20), we obtain that

$$(\forall^{\text{st}} Y^2)(\exists^{\text{st}} x^{0^*})(\forall f^1, g^1 \leq_1 1)(\exists N^0 \in x)(\overline{f}N =_0 \overline{g}N \to Y(f) =_0 Y(g)). \tag{A.21}$$

which immediately yields that

$$(\forall^{\text{st}} Y^2)(\exists^{\text{st}} N_0^0)(\forall f^1, g^1 \leq_1 1)(\overline{f}N_0 =_0 \overline{g}N_0 \to Y(f) =_0 Y(g)), \tag{A.22}$$

by taking N_0 in (A.22) to be $\max_{i<|x|} x(i)$ for x as in (A.21). Now that $\overline{f}N_0 * 00\ldots$ is standard for *any* f^1 and standard N_0 by the basic axioms of P (See Definition A.3), Hence, $Y(\overline{f}N_0 * 00\ldots)$ is also standard and (A.22) becomes

$$(\forall^{\text{st}} Y^2)(\forall f^1 \leq_1 1)(\exists^{\text{st}} M_0^0)(Y(f) \leq M_0), \tag{A.23}$$

if we take $M_0 = Y(\overline{f}N_0 * 00\ldots)$. Applying *Idealisation* to (A.23), we obtain

$$(\forall^{\text{st}} Y^2)(\exists^{\text{st}} y^{0^*})(\forall f^1 \leq_1 1)(\exists M_0^0 \in y)(Y(f) \leq M_0), \tag{A.24}$$

and defining N as the maximum of all elements in y in (A.24), we obtain

$$(\forall^{\text{st}} Y^2)(\exists^{\text{st}} N^0)(\forall f^1 \leq_1 1)(Y(f) \leq N), \tag{A.25}$$

Now fix some standard g^2 in (A.14) and let N_1 be its standard upper bound on Cantor space from (A.25). Define the required (standard) w as follows: $w(2)$ is N_1 and $w(1)$ is the finite sequence consisting of all binary sequences $\alpha_i^1 = \sigma_i^{0^*} * 00\ldots$ for $i \leq 2^{N_1}$ and $|\sigma_i| = N_1$. Then STP follows from Theorem A.14. $\square$

Corollary A.16. *From the proof in* P *that* $\mathsf{NUC} \to \mathsf{STP}$, *a term* $t^{3\to 3}$ *can be extracted such that* $\mathsf{E\text{-}PA}^{\omega *} + \mathsf{QF\text{-}AC}^{1,0}$ *proves* $(\forall \Omega^3)\big[\mathsf{MUC}(\Omega) \to \mathsf{SCF}(t(\Omega))\big]$.

Proof. The proof amounts to nothing more than applying Remark 4.8. Indeed, by the proof of the theorem, NUC is equivalent to the normal form (A.22), while (A.14) is a normal form for STP and we abbreviate the latter normal form as $(\forall^{\text{st}} g^2)(\exists^{\text{st}} w^{1^*})B(g,w)$. Hence, $\mathsf{NUC} \to \mathsf{STP}$ becomes $(\forall^{\text{st}} Y^2)(\exists^{\text{st}} N^0)A(Y,N) \to (\forall^{\text{st}} g^2)(\exists^{\text{st}} w^{1^*})B(g,w)$, which yields

$$\big[(\exists^{\text{st}} \Omega^3)(\forall Y^2)A(Y,\Omega(Y)) \to (\forall^{\text{st}} g^2)(\exists^{\text{st}} w^{1^*})B(g,w)\big], \tag{A.26}$$

by strengthening the antecedent. Bringing all standard quantifiers up front:

$$(\forall^{\rm st}\Omega^3, g^2)(\exists^{\rm st} w^{1^*})\big[(\forall Y^2)A(Y,\Omega(Y)) \to B(g,w)\big]; \tag{A.27}$$

Applying Corollary A.6 to '$\mathsf{P} \vdash$ (A.27)', we obtain a term $t^{3\to 3}$ such that

$$(\forall \Omega^3, g^2)(\exists w \in t(\Omega, g))\big[(\forall Y^2)A(Y,\Omega(Y)) \to B(g,w)\big] \tag{A.28}$$

is provable in $\mathsf{E\text{-}PA}^{\omega*} + \mathsf{QF\text{-}AC}^{1,0}$. Bringing all quantifiers inside again, (A.28) yields

$$(\forall \Omega^3)\big[(\forall Y^2)A(Y,\Omega(Y)) \to (\forall g^2)(\exists w \in t(\Omega, g))B(g,w)\big], \tag{A.29}$$

Clearly, the antecedent of (A.29) expresses that Ω^3 is the fan functional. To define a functional Θ as in $\mathsf{SCF}(\Theta)$ from $t(\Omega, g)$, note that the latter is a finite sequence of numbers and binary sequences by (A.14); Using basic sequence coding, we may assume that $t(\Omega, g) = t_0(\Omega, g) * t_1(\Omega, g)$, where the first (resp. second) part contains the binary sequences (resp. numbers). Now define $\Theta(g)(1)$ as $\max_{i<|t_1(\Omega,g)|} t_1(\Omega, g)(i)$ and $\Theta(g)(2) := t_0(\Omega, g)$, and note that Θ indeed satisfies $\mathsf{SCF}(\Theta)$. $\square$

References

[1] *The Agda proof assistant, version 2.5.1.1*, Main development website (Sept. 2016). `http://wiki.portal.chalmers.se/agda/pmwiki.php`.

[2] Sergio Albeverio, Raphael Høegh-Krohn, Jens Erik Fenstad, and Tom Lindstrøm, *Nonstandard methods in stochastic analysis and mathematical physics*, Pure and Applied Mathematics, vol. 122, Academic Press, 1986.

[3] Michèle Artigue, *Analysis*, Advanced Mathematical Thinking (ed. David O. Tall), Springer (1994).

[4] Mark van Atten and Dirk van Dalen, *Arguments for the Continuity Principle*, Bulletin of Symbolic Logic **8** (2002), 329–347.

[5] Jeremy Avigad and Solomon Feferman, *Gödel's functional ("Dialectica") interpretation*, Handbook of proof theory, Stud. Logic Found. Math., vol. 137, 1998, pp. 337–405.

[6] Andrej Bauer, *Five Stages Of Accepting Constructive Mathematics*, Bulletin (New Series) Of The American Mathematical Society (2016).

[7] Michael J. Beeson, *Foundations of constructive mathematics*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 6, Springer, 1985. Metamathematical studies.

[8] Benno van den Berg, Eyvind Briseid, and Pavol Safarik, *The strength of countable saturation*, arXiv link: `https://arxiv.org/abs/1605.02534` **163** (2016).

[9] ———, *A functional interpretation for nonstandard arithmetic*, Ann. Pure Appl. Logic **163** (2012), no. 12, 1962–1994.

[10] Ulrich Berger, *Uniform Heyting arithmetic*, Ann. Pure Appl. Logic **133** (2005), 125–148.

[11] Errett Bishop, *Foundations of constructive analysis*, McGraw-Hill Book Co., New York, 1967.

[12] ———, *Aspects of constructivism*, Notes on the lectures delivered at the Tenth Holiday Mathematics Symposium, New Mexico State University, Las Cruces, December 27-31, 1972.

[13] ———, *Review of* [53], Bull. Amer. Math. Soc **81** (1977), no. 2, 205-208.

[14] ———, *The crisis in contemporary mathematics*, Proceedings of the American Academy Workshop on the Evolution of Modern Mathematics, 1975, pp. 507–517.

[15] ———, *Mathematics as a numerical language*, Intuitionism and Proof Theory (Proc. Conf., Buffalo, N.Y., 1968), North-Holland, 1970, pp. 53–71.

[16] Douglas S. Bridges and Luminiţa Simona Vîţă, *Techniques of constructive analysis*, Universitext, Springer, New York, 2006.

[17] Douglas Bridges and Fred Richman, *Varieties of constructive mathematics*, London Mathematical Society Lecture Note Series, vol. 97, Cambridge University Press, Cambridge, 1987.

[18] L. E. J. Brouwer, *Collected works. Vol. 1*, North-Holland Publishing Co., Amsterdam, 1975. Philosophy and foundations of mathematics; Edited by A. Heyting.

[19] ———, *Collected works, Vol. 2*, North-Holland; Elsevier, 1976. Geometry, analysis, topology and mechanics; Edited by Hans Freudenthal.

[20] ———, *Over de grondslagen van de wiskunde*, University of Amsterdam, 1907. Doctoral dissertation in Dutch with English translation in [18].

[21] L.E.J. Brouwer, *Intuitionistische Betrachtungen über den Formalismus*, Sitzungsberichte der Preuszischen Akademie der Wissenschaften zu Berlin (1928), 48–52. English translation of §1 in [101], p. 490-492.

[22] L. E. J. Brouwer, *Brouwer's Cambridge lectures on intuitionism*, Cambridge University Press, 1981. D. van Dalen (ed.)

[23] Samuel R. Buss, *An introduction to proof theory*, Handbook of proof theory, Stud. Logic Found. Math., vol. 137, North-Holland, Amsterdam, 1998, pp. 1–78.

[24] Alain Connes, *Non-standard stuff*, Alain Connes' blog (2007). `http://noncommutativegeometry.blogspot.com/2007/07/non-standard-stuff.html`.

[25] ———, *An interview with Alain Connes, Part I*, EMS Newsletter **63** (2007), 25-30. `http://www.mathematics-in-europe.eu/maths-as-a-profession/interviews`.

[26] Joseph W. Dauben, *Arguments, logic and proof: mathematics, logic and the infinite*, History of mathematics and education: ideas and experiences (Essen, 1992), Stud. Wiss. Soz. Bildungsgesch. Math., vol. 11, 1996, pp. 113–148.

[27] M. Davis and M. Hausner, *Book review. The Joy of Infinitesimals. J. Keisler's Elementary Calculus*, Mathematical Intelligencer **1** (1978), 168–170.

[28] Radu Diaconescu, *Axiom of choice and complementation*, Proc. AMS **51** (1975), 176–178.

[29] Damir D. Dzhafarov, *Reverse Mathematics Zoo*. `http://rmzoo.uconn.edu/`.

[30] Philip Ehrlich, *The rise of non-Archimedean mathematics and the roots of a misconception. I. The emergence of non-Archimedean systems of magnitudes*, Arch. Hist. Exact Sci. **60** (2006), no. 1, 1–121.

[31] Martin Escardo and Paulo Oliva, *The Herbrand Functional Interpretation of the Double Negation Shift* (2015). See `http://arxiv.org/abs/1410.4353`.

[32] Howard W. Eves, *Mathematical circles squared; a third collection of mathematical stories and anecdotes*, Eves Series in Mathematics, Prindle, Weber & Schmidt, 1972.

[33] Harvey Friedman, *Some systems of second order arithmetic and their use*, Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 1, 1975, pp. 235–242.

[34] ———, *Systems of second order arithmetic with restricted induction, I & II (Abstracts)*, Journal of Symbolic Logic **41** (1976), 557–559.

[35] ———, *Computational Nonstandard Analysis*, FOM mailing list (Sept. 1st, 2015). `https://www.cs.nyu.edu/pipermail/fom/2015-September/018984.html`.

[36] Valerii Glivenko, *On the logic of Brouwer*, Istor.-Mat. Issled. (2) **5(40)** (2000) (Russian, with English summary). Publication, comments, and translation from the French by M. M. Novosëlov.

[37] Kurt Gödel, *Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes*, Dialectica **12** (1958), 280–287 (German, with English summary).

[38] Amar Hadzihasanovic and Benno van den Berg, *Nonstandard functional interpretations and models*, To appear in Notre Dame Journal for Formal Logic (2016).

[39] Paul R. Halmos, *I want to be a mathematician*, Springer-Verlag, 1985. An automathography.

[40] Jacques Harthong and Georges Reeb, *Intuitionnisme 84*, La mathématique non standard, Fondem. Sci., CNRS, Paris, 1989, pp. 213–273 (French). With comments by Harthong.

[41] A. Heyting, *Intuitionism: An introduction*, Second revised edition, North-Holland, 1966.

[42] Arend Heyting, *Address to Professor A. Robinson. At the occasion of the Brouwer memorial lecture given by Prof. A.Robinson on the 26th April 1973*, Nieuw Arch. Wisk. **21** (1973), no. 3, 134–137.

[43] Jeffry L. Hirst, *Representations of reals in reverse mathematics*, Bull. Pol. Acad. Sci. Math. **55** (2007), no. 4, 303–316.

[44] W. A. Howard, *Appendix: Hereditarily majorizable function of finite type*, Metamathematical investigation of intuitionistic arithmetic and analysis, Springer, 1973, pp. 454–461. Lecture Notes in Math., Vol. 344.

[45] Albert E. Hurd and Peter A. Loeb, *An introduction to nonstandard real analysis*, Pure and Applied Mathematics, vol. 118, Academic Press Inc., Orlando, FL, 1985.

[46] Vladimir Lifschitz, *Calculable natural numbers*, Intensional mathematics, Stud. Logic Found. Math., vol. 113, North-Holland, 1985, pp. 173–190.

[47] Hajime Ishihara, *Reverse mathematics in Bishop's constructive mathematics*, Philosophia Scientiae (Cahier Spécial) **6** (2006), 43-59.

[48] Vladimir Kanovei, Mikhail G. Katz, and Thomas Mormann, *Tools, objects, and chimeras: Connes on the role of hyperreals in mathematics*, Found. Sci. **18** (2013), no. 2, 259–296.

[49] Mikhail G. Katz and Eric Leichtnam, *Commuting and noncommuting infinitesimals*, Amer. Math. Monthly **120** (2013), no. 7, 631–641.

[50] Karin U. Katz and Mikhail G. Katz, *Meaning in Classical Mathematics: Is it at Odds with Intuitionism*, Intellectica **56** (2011), no. 2, 223–302.

[51] Richard Kaye, *Models of Peano arithmetic*, Oxford Logic Guides, vol. 15, The Clarendon Press, 1991.

[52] H. Jerome Keisler, *The hyperreal line*, Real numbers, generalizations of the reals, and theories of continua, Synthese Lib., vol. 242, Kluwer Acad. Publ., Dordrecht, 1994, pp. 207–237.

[53] ———, *Elementary Calculus*, Prindle, Weber and Schmidt, Boston, 1976.

[54] Ulrich Kohlenbach, *Applied proof theory: proof interpretations and their use in mathematics*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2008.

[55] ———, *Higher order reverse mathematics*, Reverse mathematics 2001, Lect. Notes Log., vol. 21, ASL, 2005, pp. 281–295.

[56] ———, *Foundational and mathematical uses of higher types*, Reflections on the foundations of mathematics (Stanford, CA, 1998), Lect. Notes Log., vol. 15, ASL, 2002, pp. 92–116.

[57] ———, *On uniform weak König's lemma*, Ann. Pure Appl. Logic **114** (2002), no. 1-3, 103–116. Commemorative Symposium Dedicated to Anne S. Troelstra (Noordwijkerhout, 1999).

[58] Georg Kreisel, *Mathematical logic*, Lectures on Modern Mathematics, Vol. III, Wiley, New York, 1965, pp. 95–195.

[59] John Longley and Dag Normann, *Higher-order Computability*, Theory and Applications of Computability, Springer, 2015.

[60] Per Martin-Löf, *An intuitionistic theory of types: predicative part*, Logic Colloquium '73 (Bristol, 1973), North-Holland, Amsterdam, 1975, pp. 73–118. Studies in Logic and the Foundations of Mathematics, Vol. 80.

[61] Ieke Moerdijk and Erik Palmgren, *Minimal models of Heyting arithmetic*, J. Symbolic Logic **62** (1997), no. 4, 1448–1460.

[62] Edward Nelson, *Internal set theory: a new approach to nonstandard analysis*, Bull. Amer. Math. Soc. **83** (1977), no. 6, 1165–1198.

[63] ———, *Unfinished book on nonstandard Analysis* (2016), 1–26. Chapter 1 from `https://web.math.princeton.edu/~nelson/books.html`.

[64] ———, *Predicative arithmetic*, Mathematical Notes, vol. 32, Princeton University Press, Princeton, NJ, 1986.

[65] Dag Normann and Sam Sanders, *Nonstandard Analysis, Computability Theory, and their surprising connections*, Preprint (2016).

[66] Dag Normann, *Recursion on the countable functionals*, LNM 811, vol. 811, Springer, 1980.

[67] Horst Osswald, *Computation of the kernels of Lévy functionals and applications*, Illinois Journal of Mathematics **55** (2011), no. 3, 815–833.

[68] ———, *Malliavin calculus for Lévy processes and infinite-dimensional Brownian motion*, Cambridge Tracts in Mathematics, vol. 191, Cambridge University Press, 2012.

[69] Erik Palmgren, *A sheaf-theoretic foundation for nonstandard analysis*, Ann. Pure Appl. Logic **85** (1997), no. 1, 69–86.

[70] ———, *Developments in constructive nonstandard analysis*, B. Sym. Logic (1998), 233–272.

[71] Michael Rathjen, *The disjunction and related properties for constructive Zermelo-Fraenkel set theory*, J. Symbolic Logic **70** (2005), no. 4, 1232–1254.

[72] Fred Richman, *The frog replies*, The Mathematical Intelligencer **9** (1987), no. 3, 22–24.

[73] ——— (ed.), *Constructive mathematics*, Lecture Notes in Mathematics, vol. 873, Springer, 1981.

[74] Abraham Robinson, *Non-standard analysis*, North-Holland, Amsterdam, 1966.

[75] ———, *Formalism* 64, Logic, Methodology and Philos. Sci. (Proc. 1964 Internat. Congr.), North-Holland, 1965, pp. 228–246.

[76] David A. Ross, *The constructive content of nonstandard measure existence proofs—is there any?*, Reuniting the antipodes—constructive and nonstandard views of the continuum (Venice, 1999), Synthese Lib., vol. 306, Kluwer, 2001, pp. 229–239.

[77] Jean-Michel Salanskis, *Le constructivisme non standard*, Histoire des Sciences. [History of Science], Presses Universitaires du Septentrion, Villeneuve d'Ascq, 1999 (French, with French summary).
[78] Sam Sanders, *The Gandy-Hyland functional and a hitherto unknown computational aspect of Nonstandard Analysis*, Submitted, Available from: `http://arxiv.org/abs/1502.03622` (2015).
[79] ———, *The taming of the Reverse Mathematics zoo*, Submitted, `http://arxiv.org/abs/1412.2022` (2015).
[80] ———, *Nonstandard Analysis And Constructivism!*, In preparation (2016).
[81] ———, *The unreasonable effectiveness of Nonstandard Analysis*, Submitted to APAL special issue of LFCS, `http://arxiv.org/abs/1508.07434` (2016).
[82] ———, *Formalism16*, To appear in Synthese, Special Issue for SOTFOMIII (2017).
[83] ———, *On the connection between Nonstandard Analysis and Computability Theory* (2016). `http://arxiv.org/abs/1609.01919`.
[84] ———, *The computational content of the Loeb measure* (2016). `http://arxiv.org/abs/1609.01945`.
[85] ———, *The refining of the taming of the Reverse Mathematics zoo*, To appear in Notre Dame Journal for Formal Logic, `http://arxiv.org/abs/1602.02270` (2016).
[86] Stephen G. Simpson (ed.), *Reverse mathematics 2001*, Lecture Notes in Logic, vol. 21, ASL, La Jolla, CA, 2005.
[87] ———, *Subsystems of second order arithmetic*, 2nd ed., Perspectives in Logic, CUP, 2009.
[88] Stephen G. Simpson and Keita Yokoyama, *A nonstandard counterpart of* WWKL, Notre Dame J. Form. Log. **52** (2011), no. 3, 229–243.
[89] Helmut Schwichtenberg, *Minlog, the proof assistant*, Official website: `http://www.minlog-system.de` (2016).
[90] Ian Stewart, *Frog and Mouse Revisited*, The Mathematical Intelligencer **8** (1986), no. 4, 72-82.
[91] Gabriel Stolzenberg, *Letter to the editor*, Notices Amer. Math. Soc. **25** (1978), 242.
[92] ———, *Reply to Russel O'Conner*, Constructive News `https://groups.google.com/forum/#!forum/constructivenews` (Oct. 3, 2009).
[93] David Tall, *Natural and Formal Infinities*, Educational Studies in Mathematics **48** (2001), 2–3.
[94] Terence Tao, *Entries on Nonstandard Analysis*, 2015. Wordpress blog, `https://terrytao.wordpress.com/tag/nonstandard-analysis/`.
[95] ———, *Structure and randomness*, American Mathematical Society, Providence, RI, 2008. Pages from year one of a mathematical blog.
[96] Anne Sjerp Troelstra, *Principles of intuitionism*, Lectures presented at the Summer Conference on Intuitionism and Proof Theory (1968) at SUNY at Buffalo, N.Y. Lecture Notes in Mathematics, Vol. 95, Springer-Verlag, Berlin-New York, 1969. MR0244003
[97] ———, *Metamathematical investigation of intuitionistic arithmetic and analysis*, Springer Berlin, 1973. Lecture Notes in Mathematics, Vol. 344.
[98] Anne Sjerp Troelstra and Dirk van Dalen, *Constructivism in mathematics. Vol. I*, Studies in Logic and the Foundations of Mathematics, vol. 121, North-Holland, 1988.
[99] Dirk van Dalen, *L. E. J. Brouwer—topologist, intuitionist, philosopher*, Springer, 2013. How mathematics is rooted in life.
[100] ———, *How connected is the intuitionistic continuum?*, J. Symbolic Logic **62** (1997), no. 4, 1147–1150.
[101] Jean van Heijenoort, *From Frege to Gödel. A source book in mathematical logic, 1879–1931*, Harvard University Press, Cambridge, Mass., 1967.
[102] Walter P. van Stigt, *Brouwer's intuitionism*, Studies in the History and Philosophy of Mathematics, vol. 2, North-Holland, 1990.
[103] Martin Väth, *Nonstandard analysis*, Birkhäuser Verlag, Basel, 2007.
[104] Guy Wallet, *Integer calculus on the Harthong-Reeb line*, ARIMA Rev. Afr. Rech. Inform. Math. Appl. **9** (2008), 517–536 (English, with English and French summaries).
[105] F. Wattenberg, *Nonstandard analysis and constructivism?*, Studia Logica **47** (1988), 303–309.

[106] Freek Wiedijk, *The Seventeen Provers of the World: Foreword by Dana S. Scott (Lecture Notes in Computer Science / Lecture Notes in Artificial Intelligence)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[107] Manfred Wolff and Peter A. Loeb (eds.), *Nonstandard analysis for the working mathematician*, Mathematics and its Applications, vol. 510, Kluwer Academic Publishers, 2000.

[108] ——— (ed.), *Nonstandard analysis for the working mathematician*, Mathematics and its Applications, vol. 510, Kluwer, 2015. Second edition.

[109] Chuangjie Xu and Sam Sanders, *Extracting the computational content of Nonstandard Analysis*, In preparation; Agda code: `http://cj-xu.github.io/agda/nonstandard_dislectica/Dialectica.html` (2016).

[110] Xiaokang Yu, *Lebesgue convergence theorems and reverse mathematics*, Math. Logic Quart. **40** (1994), no. 1, 1–13.

[111] Xiaokang Yu and Stephen G. Simpson, *Measure theory and weak König's lemma*, Arch. Math. Logic **30** (1990), no. 3, 171–180.

[112] The Univalent Foundations Program, *Homotopy type theory—univalent foundations of mathematics*, Institute for Advanced Study, Princeton, 2013. `http://homotopytypetheory.org/book`.