Some nonstandard equivalences in Reverse Mathematics

Sam Sanders*

Center for Advanced Studies &
Munich Center for Mathematical Philosophy,
LMU Munich, Germany sasander@me.com

Abstract. Reverse Mathematics (RM) is a program in the foundations of mathematics founded by Friedman and developed extensively by Simpson. The aim of RM is finding the minimal axioms needed to prove a theorem of ordinary (i.e. non-set theoretical) mathematics. In the majority of cases, one also obtains an *equivalence* between the theorem and its minimal axioms. This equivalence is established in a weak logical system called *the base theory*; four prominent axioms which boast lots of such equivalences are dubbed *mathematically natural* by Simpson. In this paper, we show that a number of axioms from Nonstandard Analysis are equivalent to theorems of ordinary mathematics *not involving Nonstandard Analysis*. These equivalences are proved in a *weak* base theory recently introduced by van den Berg and the author. In particular, our base theories have the first-order strength of *elementary function arithmetic*, in contrast to the original version of this paper [22]. Our results combined with Simpson's criterion for naturalness suggest the controversial point that Nonstandard Analysis is actually mathematically natural.

1 Introduction

Reverse Mathematics (RM) is a program in the foundations of mathematics founded by Friedman ([5]) and developed extensively by Simpson ([24]) and others. We refer to the latter for an overview of RM and will assume basic familiarity, in particular with the *Big Five* systems of RM. The latter are (still) claimed to capture the majority of theorems of ordinary (i.e. non-set theoretical) mathematics ([13, p. 495]). Our starting point is the following quote by Simpson on the 'mathematical naturalness' of logical systems from [24, I.12]:

> From the above it is clear that the [Big Five] five basic systems RCA_0, WKL_0, ACA_0, ATR_0, Π^1_1-CA_0 arise naturally from investigations of the Main Question. The proof that these systems are mathematically natural is provided by Reverse Mathematics.

* This research was supported by the Alexander von Humboldt Foundation and LMU Munich (via the Excellence Initiative).

In a nutshell, according to Simpson, the many equivalences in RM, proved over RCA_0 and involving the other four Big Five, imply that the Big Five systems are mathematically natural. In this paper, we show that a number of axioms from *Nonstandard Analysis* (NSA) are equivalent to theorems of ordinary mathematics *not involving NSA*. These results combined with Simpson's criterion for naturalness suggest the controversial point that NSA is actually mathematically natural. Indeed, both Alain Connes and Errett Bishop have expressed extremely negative (but unfounded; see [20]) opinions of NSA, in particular its naturalness.

Finally, the aforementioned equivalences are proved in a (weak) base theory recently introduced by van den Berg and the author in [3]. We sketch the main properties of this base theory in Section 2 and prove our main results in Section 3. The latter include an equivalence between the *Heine-Borel compactness* (for **any** open cover) of the unit interval and the *nonstandard compactness* of Cantor space. We obtain similar results based on WWKL, a weakening of WKL. The main improvement over [22] is that the base theories in this paper are Π_2^0-conservative over *elementary function arithmetic* (See [24, II.8] for the latter).

2 A base theory from Nonstandard Analysis

We introduce the system B_0 from [3]. This system is a Π_2^0-conservative extension of EFA (aka $I\Delta_0 + \mathsf{EXP}$) enriched with all finite types and fragments of Nelson's axioms of *internal set theory* ([14]), a well-known axiomatic approach to NSA.

Let $\mathsf{E\text{-}EFA}^\omega$ be EFA enriched with all finite types, i.e. Kohlenbach's system $\mathsf{E\text{-}G_3A}^\omega$ ([9, p. 55]). The language of B_0 is obtained from that of $\mathsf{E\text{-}EFA}^\omega$ by adding unary predicates 'st^σ' for any finite type σ. Formulas in the old language of $\mathsf{E\text{-}EFA}^\omega$, i.e. those not containing these new symbols, are *internal*; By contrast, general formulas of B_0 are *external*. The new 'st' predicates give rise to two new quantifiers as in (2.1), and we omit type superscripts whenever possible.

$$(\forall^{\mathrm{st}} x)\Phi(x) \equiv (\forall x)(\,\mathrm{st}(x) \to \Phi(x)) \text{ and } (\exists^{\mathrm{st}} x)\Phi(x) \equiv (\exists x)(\,\mathrm{st}(x) \wedge \Phi(x)). \quad (2.1)$$

The system B_0 is $\mathsf{E\text{-}EFA}^\omega + \mathsf{QF\text{-}AC}^{1,0}$, plus the basic axioms as in Definition 2.2, and fragments of Nelson's axioms of internal set theory IST, namely *Idealisation* I, *Standardisation* $\mathsf{HAC}_{\mathsf{int}}$, and *Transfer* $\mathsf{PF\text{-}TP}_\forall$, defined as follows.

Definition 2.1 [$\mathsf{QF\text{-}AC}$] For all finite types σ, τ and quantifier-free A:

$$(\forall x^\sigma)(\exists y^\tau)A(x,y) \to (\exists Y^{\sigma\to\tau})(\forall x^\sigma)A(x,Y(x)) \qquad (\mathsf{QF\text{-}AC}^{\sigma,\tau})$$

Definition 2.2

1. *The axioms* $\mathrm{st}(x) \wedge x = y \to \mathrm{st}(y)$ *and* $\mathrm{st}(f) \wedge \mathrm{st}(x) \to \mathrm{st}(fx)$.
2. *The axiom* $\mathrm{st}(t)$ *for each term* t *in the language of* B_0.
3. *The axiom* $\mathrm{st}_0(x) \wedge y \leq_0 x \to \mathrm{st}_0(y)$.

For the next definition, we note that x in (2.3) and $F(x)$ in (2.2) are both a *finite sequence* of objects of type σ, as discussed in Notation 2.4.

Definition 2.3 [Fragments of IST]

1. $\mathsf{HAC}_{\mathsf{int}}$: For any internal formula φ, we have
$$(\forall^{\mathrm{st}} x^{\rho})(\exists^{\mathrm{st}} y^{\sigma})\varphi(x,y) \to \big(\exists^{\mathrm{st}} F^{\rho\to\sigma}\big)(\forall^{\mathrm{st}} x^{\rho})(\exists y^{\sigma} \in F(x))\varphi(x,y). \tag{2.2}$$
2. I: For any internal formula φ, we have
$$(\forall^{\mathrm{st}} x^{\sigma})(\exists y^{\tau})(\forall z^{\sigma} \in x)\varphi(z,y) \to (\exists y^{\tau})(\forall^{\mathrm{st}} z^{\sigma})\varphi(z,y). \tag{2.3}$$
3. $\mathsf{PF\text{-}TP}_{\forall}$: For any internal φ with all parameters shown, we have $\forall^{\mathrm{st}} x\,\varphi(x) \to \forall x\,\varphi(x)$, i.e. x is the **only** free variable in $\varphi(x)$.

Notation 2.4 (Finite sequences) There are at least two ways of approaching 'finite sequences of objects of type σ' in $\mathsf{E\text{-}EFA}^{\omega}$: First of all, as in [2], we could extend $\mathsf{E\text{-}EFA}^{\omega}$ with types σ^* for finite sequences of objects of type σ, add constants for the empty sequence and the operation of prepending an element to a sequence, as well as a list recursor satisfying the expected equations.

Secondly, as in [3], we could exploit the fact that one can code finite sequences of objects of type σ as a single object of type σ in such a way that every object of type σ codes a sequence. Moreover, the operations on sequences, such as extracting their length or concatenating them, are given by terms in Gödel's T.

We choose the second option here and will often use set-theoretic notation as follows: '$\emptyset$' is (the code of) the empty sequence,' $\cup$' stands for concatenation, and '$\{x\}$' for the finite sequence of length 1 with sole component x. For x and y of the same type we will write $x \in y$ if x is equal to one of the components of the sequence coded by y. Furthermore, for $\alpha^{0\to\rho}$ and k^0, the finite sequence $\overline{\alpha}k$ is exactly $\langle \alpha(0), \alpha(1), \dots, \alpha(k-1)\rangle$. Finally, if Y is of type $\sigma \to \tau$ and x is of type σ we define $Y[x]$ of type τ as $Y[x] := \cup_{f\in Y} f(x)$.

With this notation in place, we can now formulate a crucial theorem from [3, §3].

Theorem 2.5 *For φ internal and Δ_{int} a collection of internal formulas, if the system $\mathsf{B}_0 + \Delta_{\mathsf{int}}$ proves $(\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\varphi(x,y)$, then*
$$\mathsf{E\text{-}EFA}^{\omega} + \mathsf{QF\text{-}AC}^{1,0} + \Delta_{\mathsf{int}} \vdash (\exists \Phi)(\forall x)(\exists y \in \Phi(x))\varphi(x,y) \tag{2.4}$$

By the results in Section 3 and [3], $\mathsf{PF\text{-}TP}_{\forall}$ is useful for obtaining equivalences as in RM. However, this 'usefulness' comes at a price, as B_0^- (i.e. $\mathsf{B}_0 \setminus \mathsf{PF\text{-}TP}_{\forall}$) satisfies the following, where a *term of Gödel's system T is obtained*, to be compared to the *existence* of a functional in (2.4). Hence, $\mathsf{PF\text{-}TP}_{\forall}$ seems usuitable for proof mining, as the latter deals with extracted *terms*

Theorem 2.6 (Term extraction) *If Δ_{int} is a collection of internal formulas and ψ is internal, and $\mathsf{B}_0^- + \Delta_{\mathsf{int}} \vdash (\forall^{\mathrm{st}} x)(\exists^{\mathrm{st}} y)\psi(x,y)$, then one can extract from the proof a term t from Gödel's T such that*
$$\mathsf{E\text{-}EFA}^{\omega} + \mathsf{QF\text{-}AC}^{1,0} + \Delta_{\mathsf{int}} \vdash (\forall x)(\exists y \in t(x))\psi(x,y). \tag{2.5}$$

We finish this section with some notations.

Notation 2.7 (Equality) The system $\mathsf{E\text{-}EFA}^\omega$ includes equality '$=_0$' for numbers as a primitive. Equality '$=_\tau$' for x^τ, y^τ is:

$$[x =_\tau y] \equiv (\forall z_1^{\tau_1} \dots z_k^{\tau_k})[xz_1 \dots z_k =_0 yz_1 \dots z_k], \tag{2.6}$$

if the type τ is composed as $\tau \equiv (\tau_1 \to \dots \to \tau_k \to 0)$. Inequality '$\leq_\tau$' is (2.6) with $=_0$ replaced by $\leq_0$. Similarly, we define 'approximate equality $\approx_\tau$' as:

$$[x \approx_\tau y] \equiv (\forall^{\text{st}} z_1^{\tau_1} \dots z_k^{\tau_k})[xz_1 \dots z_k =_0 yz_1 \dots z_k] \tag{2.7}$$

Notation 2.8 (Real numbers and related notions in B_0)

1. Natural numbers correspond to type zero objects. Rational numbers are defined 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. $(\forall n^0, i^0)(|q_n - q_{n+i})| <_0 \frac{1}{2^n})$.
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 $(\forall n^0)(|q_n - r_n| \leq \frac{1}{2^{n-1}})$. 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 $F : \mathbb{R} \to \mathbb{R}$ are represented by $\Phi^{1\to 1}$ such that
$$(\forall x, y)(x =_\mathbb{R} y \to \Phi(x) =_\mathbb{R} \Phi(y)). \tag{RE}$$
7. Sets of natural numbers $X^1, Y^1, Z^1, \dots$ are represented by binary sequences.

Notation 2.9 (Using $\mathsf{HAC}_{\mathsf{int}}$) As noted in Notation 2.4, finite sequences play an important role in B_0. In particular, $\mathsf{HAC}_{\mathsf{int}}$ produces a functional which outputs a *finite sequence* of witnesses. However, $\mathsf{HAC}_{\mathsf{int}}$ provides an actual *witnessing functional* assuming (i) $\tau = 0$ in $\mathsf{HAC}_{\mathsf{int}}$ and (ii) the formula φ from $\mathsf{HAC}_{\mathsf{int}}$ is 'sufficiently monotone' as in: $(\forall^{\text{st}} x^\sigma, n^0, m^0)\big([n \leq_0 m \wedge \varphi(x, n)] \to \varphi(x, m)\big)$. Indeed, in this case one simply defines $G^{\sigma+1}$ by $G(x^\sigma) := \max_{i<|F(x)|} F(x)(i)$ which satisfies $(\forall^{\text{st}} x^\sigma)\varphi(x, G(x))$. To save space in proofs, we will sometimes skip the (obvious) step involving the maximum of finite sequences, when applying $\mathsf{HAC}_{\mathsf{int}}$. We assume the same convention for other finite sequences e.g. obtained from Theorem 2.6, or the contraposition of idealisation I.

3 Reverse Mathematics and Nonstandard Analysis

In Sections 3.1 and 3.2, we establish the equivalence between the *nonstandard compactness* of Cantor space and the *Heine-Borel compactness* (for **any** open cover) of the unit interval. The latter essentially predates[1] set theory, and is hence definitely part of 'ordinary mathematics' in the sense of RM. We establish similar results for theorems based on WWKL in Section 3.3. We shall use 'computable' in the sense of Kleene's schemes S1-S9 inside ZFC ([12, §5.1.1]).

[1] Heine-Borel compactness was studied before 1895 by Cousin ([4, p. 22]). The collected works of Pincherle contain a footnote by the editors ([18, p. 67]) stating that the associated *Teorema* (from 1882) corresponds to the Heine-Borel theorem.

3.1 Nonstandard compactness and the special fan functional

The main result of this section is Theorem 3.4, which establishes an equivalence involving the nonstandard compactness of Cantor space and the *special fan functional*, introduced in [19] and studied in detail in [15]. The variable 'T' is reserved for trees, and '$T \leq_1 1$' means that T is a binary tree.

Definition 3.1 [Special fan functional] We define $\mathsf{SCF}(\Theta)$ as follows for $\Theta^{(2\to(0\times 1))}$:

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

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

From a computability theoretic perspective, the main property of Θ is the selection of $\Theta(g)(2)$ as a finite sequence of binary sequences $\langle f_0, \dots, f_n\rangle$ such that the neighbourhoods defined from $\overline{f_i}g(f_i)$ for $i \leq n$ form a cover of Cantor space; almost as a by-product, $\Theta(g)(1)$ can then be chosen to be the maximal value of $g(f_i)+1$ for $i \leq n$. No type two functional computes Θ such that $\mathsf{SCF}(\Theta)$ ([15]), while the following functional can compute Θ via a term of Gödel's T ([21]).

$$(\exists \xi^3)(\forall Y^2)\big[(\exists f^1)(Y(f)=0) \leftrightarrow \xi(Y)=0\big]. \qquad (\exists^3)$$

We stress that g^2 in $\mathsf{SCF}(\Theta)$ may be *discontinuous* and that Kohlenbach has argued for the study of discontinuous functionals in higher-order RM ([10]). Furthermore, $\mathsf{RCA}_0^\omega + (\exists\Theta)\mathsf{SCF}(\Theta)$ is conservative over WKL_0 ([15, 19]), and Θ naturally emerges from Tao's notion of *metastability*, as discussed in [16, 21, 23].

The special fan functional arose from STP, the *nonstandard compactness* of Cantor space as in *Robinson's theorem* ([7]). This fragment of *Standard Part* is also known as the 'nonstandard counterpart of weak König's lemma' ([8]).

$$(\forall \alpha^1 \leq_1 1)(\exists^{\text{st}} \beta^1 \leq_1 1)(\alpha \approx_1 \beta), \qquad (\mathsf{STP})$$

as explained by the equivalence between STP and (3.2), as follows.

Theorem 3.2 *In* B_0^-, STP *is equivalent to the following:*

$$(\forall^{\text{st}} g^2)(\exists^{\text{st}} w^1 \leq_1 1, k^0)\big[(\forall T^1 \leq_1 1)\big((\forall \alpha^1 \in w)(\overline{\alpha}g(\alpha) \not\in T) \to (\forall \beta \leq_1 1)(\exists i \leq k)(\overline{\beta}i \not\in T)\big)\big], \qquad (3.1)$$

as well as to the following:

$$(\forall T^1 \leq_1 1)\big[(\forall^{\text{st}} n)(\exists \beta)(|\beta| = n \wedge \beta \in T) \to (\exists^{\text{st}} \alpha^1 \leq_1 1)(\forall^{\text{st}} n)(\overline{\alpha}n \in T)\big]. \qquad (3.2)$$

Furthermore, B_0^- *proves* $(\exists^{\text{st}}\Theta)\mathsf{SCF}(\Theta) \to \mathsf{STP}$.

Proof. A detailed proof may be found in any of the following: [15, 19, 23]. In a nutshell, the implication (3.1)$\leftarrow$(3.2) follows by taking the contraposition of the latter and introducing standard g^2 in the antecedent of the resulting formula. One then uses *Idealisation* I to pull the standard quantifiers to the front and obtains (3.1). The other implication follows by pushing the standard quantifiers in the latter back inside. For the remaining implication $\mathsf{STP} \to$ (3.2) (the other one and the final part then being trivial), one uses *overspill* (See [2, §3]) to obtain a sequence of nonstandard length for a tree $T \leq_1 1$ satisfying the antecedent of (3.2), and STP converts this sequence into a standard path in T. □

For the below results, we need the following corollary which expresses the (trivial but important) fact that the type of the universal quantifier in STP (and equivalent formulations) may be lowered. We view $\alpha^0 \leq_0 1$ as a finite binary sequence; we define $\hat{\alpha}$ to be $\alpha * 00\ldots$, i.e. the type one object obtained by concatenating α with 0^1. Similarly, $T^0 \leq_0 1$ is a binary tree of type zero, and $\mathsf{SCF}_0(\Theta)$ is the specification of Θ restricted to trees $T^0 \leq_0 1$.

Corollary 3.3 *In* B_0^-, STP *is equivalent to* $(\forall \alpha^0 \leq_0 1)(\exists^{\text{st}} \beta^1 \leq 1)(\hat{\alpha} \approx_1 \beta)$, *and also to the following:*

$$(\forall T^0 \leq_0 1)\big[(\forall^{\text{st}} n)(\exists \beta)(|\beta| = n \wedge \beta \in T) \to (\exists^{\text{st}} \alpha^1 \leq_1 1)(\forall^{\text{st}} n)(\overline{\alpha} n \in T)\big], \quad (3.3)$$

and also the following:

$$(\forall^{\text{st}} g^2)(\exists^{\text{st}} w^1 \leq_1 1, k^0)\big[(\forall T^0 \leq_0 1)\big((\forall \alpha^1 \in w)(\overline{\alpha} g(\alpha) \notin T) \to (\forall \beta \leq_1 1)(\exists i \leq k)(\overline{\beta} i \notin T)\big)\big]. \quad (3.4)$$

The system $\mathsf{E\text{-}EFA}^\omega + \mathsf{QF\text{-}AC}^{1,0}$ *proves* $(\exists \Theta)\mathsf{SCF}(\Theta) \leftrightarrow (\exists \Theta_0)\mathsf{SCF}_0(\Theta_0)$.

Proof. Now, (3.3) $\leftrightarrow$ (3.4) follows in the same way as for (3.2) $\leftrightarrow$ (3.1). The first forward implication is trivial while the first reverse implication follows by considering $\overline{\alpha} N$ for nonstandard N^0 and $\alpha^1 \leq_1 1$. The implication $\mathsf{STP} \to$ (3.3) follows in the same way as in the proof of the theorem. Note that $T^0 \leq_0 1$ as in the antecedent of (3.3) must be nonstandard by the basic axioms in Definition 2.2. The implication (3.3) $\to$ (3.2) follows by restricting T^1 to sequences of some fixed nonstandard length, which yields a type zero object. The final equivalence follows by applying Theorem 2.6 to '$\mathsf{B}_0^- \vdash (3.1) \leftrightarrow (3.4)$'. □

The following theorem was proved in [3, 22] using respectively the *Suslin functional* and *Turing jump functional* $\exists^2$, rather than the *much weaker* fan functional (FF), where '$Y^2 \in \mathsf{cont}$' means that Y is continuous on $\mathbb{N}^{\mathbb{N}}$.

$$(\exists \Phi^3)(\forall Y^2 \in \mathsf{cont})(\forall f, g \leq 1)(\overline{f}\Phi(Y) = \overline{g}\Phi(Y) \to Y(f) = Y(g)) \qquad (\mathsf{FF})$$

$$(\exists \varphi^2)(\forall f^1)\big[(\exists n)(f(n) = 0) \leftrightarrow \varphi(f) = 0\big]. \qquad (\exists^2)$$

Note that the base theory for the equivalence is conservative over WKL_0^*, i.e. the first-order strength is that of elementary function arithmetic. Another noteworthy fact is that STP deals with second-order objects, while Θ is fourth-order.

Theorem 3.4 *The system* $\mathsf{B}_0 + (\mathsf{FF}) + \mathsf{QF\text{-}AC}^{2,1}$ *proves* $\mathsf{STP} \leftrightarrow (\exists \Theta)\mathsf{SCF}(\Theta)$, *while the system* $\mathsf{B}_0^- + (\exists^3) + \mathsf{QF\text{-}AC}$ *does not.*

Proof. The reverse implication is immediate using $\textsf{PF-TP}_\forall$ and Theorem 3.2. We now prove the forward implication in $\textsf{B}_0 + (\textsf{FF}) + \textsf{QF-AC}^{2,1}$, first additionally assuming $(\exists^2)$ and then again additionally assuming $\neg(\exists^2)$. The law of excluded middle then yields this implication over $\textsf{B}_0 + (\textsf{FF}) + \textsf{QF-AC}^{2,1}$. Hence, assume $(\exists^2)$ and note that $\textsf{STP}$ implies (3.4) by Corollary 3.3. Drop the second 'st' in (3.4), and apply $\textsf{PF-TP}_\forall$ to the resulting formula to obtain

$$(\forall g^2)(\exists w^1 \leq_1 1, k^0)\big[(\forall T^0 \leq_0 1)[(\forall \alpha^1 \in w)(\overline{\alpha}g(\alpha) \notin T) \\ \rightarrow (\forall \beta \leq_1 1)(\exists i \leq k)(\overline{\beta}i \notin T)]\big], \tag{3.5}$$

where the formula in big square brackets is equivalent to a quantifier-free one, thanks to $(\exists^2)$. Apply $\textsf{QF-AC}^{2,1}$ to (3.5) to obtain Θ_0 producing w^1, k^0 from g^2 as in (3.5). Corollary 3.3 yield $(\exists\Theta)\textsf{SCF}(\Theta)$. Again for the forward implication, assume $\neg(\exists^2)$ and note that all functions on $\mathbb{N}^\mathbb{N}$ are continuous by [10, Cor. 3.7]. Hence, the fan functional Φ as in $(\textsf{FF})$ applies to all functions on $\mathbb{N}^\mathbb{N}$, and we may define $\Theta(g)(2)$ as consisting of all $2^{\Phi(g)}$ binary sequences $\sigma * 00\ldots$ where $|\sigma| = \Phi(g)$ and $\Theta(g)(1) := \Phi(g)$. The forward implication now follows.

The non-implication follows from [15, Theorem 4.2] as the latter expresses that the special fan functional is not computable in any type two functional. Indeed, $\textsf{STP}$ is equivalent to (3.1) by Theorem 3.2 and applying Theorem 2.6 to $\textsf{B}_0^- + (\exists^3) + \textsf{QF-AC} + (\exists\Theta)\textsf{SCF}(\Theta) \vdash (3.1)$, one obtains a term t of Gödel's T such that $\textsf{SCF}(t)$, which is impossible. □

3.2 Nonstandard compactness and Heine-Borel compactness

We prove an equivalence between $\textsf{STP}$ and the Heine-Borel theorem *in the general*[2] *case*, i.e. the statement that any (possibly uncountable) open cover of the unit interval has a finite sub-cover. In particular, any $\Psi : \mathbb{R} \rightarrow \mathbb{R}^+$ gives rise to a 'canonical' open cover $\cup_{x\in[0,1]} I_x$ of $[0,1]$ where $I_x^\Psi \equiv (x - \Psi(x), x + \Psi(x))$. Hence, the Heine-Borel theorem trivially implies the following statement:

$$(\forall \Psi : \mathbb{R} \rightarrow \mathbb{R}^+)(\exists w^1)(\forall x \in [0,1])(\exists y \in w)(x \in I_y^\Psi). \tag{HBU}$$

By Footnote 1, $\textsf{HBU}$ is part of ordinary mathematics as it predates set theory. Furthermore, $\textsf{HBU}$ is equivalent to many basic properties of the *gauge integral* ([17]). The latter is an extension of Lebesgue's integral and provides a (direct) formalisation of the Feyman path integral. Note that the following theorem was proved using $(\exists^2)$ in the base theory in [22].

Theorem 3.5 *The system* $\textsf{B}_0 + \textsf{QF-AC}^{2,1}$ *proves that* $\textsf{STP} \leftrightarrow \textsf{HBU} \leftrightarrow \textsf{HBU}^{\text{st}}$*, while the system* $\textsf{B}_0^- + (\exists^3) + \textsf{QF-AC}$ *does not.*

Proof. For the first reverse implication, we use the same 'excluded middle trick' as in the proof of Theorem 3.4. Hence, assuming $\neg(\exists^2)$, all functionals on $\mathbb{R}$ are continuous, and $\textsf{STP} \rightarrow \textsf{WKL}^{\text{st}} \rightarrow \textsf{WKL}$ (See [3, Theorem 3.7]) yields that all

[2] The Heine-Borel theorem in RM is restricted to *countable* covers ([24, IV.1]).

functionals on $[0,1]$ are uniformly continuous by [11, §4], and HBU is immediate. Next, assuming $(\exists^2)$, note the latter allows us to (uniformly) convert reals into their binary representation (choosing the one with trailing zeros in case of non-uniqueness). Hence, any type two functional can be modified to satisfy (RE) from Notation 2.8 if necessary. Hence, HBU immediately generalises to *any* Ψ^2. Now, for the reverse implication, note that HBU trivially implies

$$(\forall \Psi^2)(\exists w^1)\underline{(\forall q^0 \in [0,1])(\exists y \in w)(|q-y| < \tfrac{1}{\Psi(y)+1})}, \tag{3.6}$$

where the underlined formula in (3.6) may be treated as quantifier-free, due to the presence of $(\exists^2)$ in the base theory. Applying $\mathsf{QF\text{-}AC}$ to (3.6), we obtain:

$$(\exists \Phi^{2\to 1})(\forall \Psi^2)(\forall q^0 \in [0,1])(\exists y \in \Phi(\Psi))(|q-y| < \tfrac{1}{\Psi(y)+1}), \tag{3.7}$$

and applying $\mathsf{PF\text{-}TP}_\forall$ to (3.7) implies that

$$(\exists^{\mathrm{st}} \Phi^{2\to 1})(\forall \Psi^2)(\forall q^0 \in [0,1])(\exists y \in \Phi(\Psi))(|q-y| < \tfrac{1}{\Psi(y)+1}), \tag{3.8}$$

and we now show that (3.8) implies STP. Since standard functionals yield standard outputs for standard inputs by Definition 2.2, (3.8) immediately implies

$$(\forall^{\mathrm{st}} \Psi^2)(\forall q^0 \in [0,1])(\exists^{\mathrm{st}} y^1 \in [0,1])(|q-y| < \tfrac{1}{\Psi(y)+1}).$$

Now, $(\forall^{\mathrm{st}}\Psi^2)(\exists^{\mathrm{st}} y \in [0,1])(|q-y| < \frac{1}{\Psi(y)+1})$ implies $(\exists^{\mathrm{st}} y \in [0,1])(q \approx y)$; indeed, $(\forall^{\mathrm{st}} y \in [0,1])(q \not\approx y)$ implies $(\forall^{\mathrm{st}} y \in [0,1])(\exists^{\mathrm{st}} k^0)(|q-y| \geq \frac{1}{k})$, and applying $\mathsf{HAC}_{\mathsf{int}}$ yields standard Ξ^2 such that $(\forall^{\mathrm{st}} y \in [0,1])(\exists k^0 \in \Xi(y))(|q-y| \geq \frac{1}{k})$. Defining standard Ψ_0^2 as $\Psi_0(y) := \max_{i<|\Xi(y)|} \Xi(y)(i)$, we obtain $(\forall^{\mathrm{st}} y \in [0,1]))(|q-y| \geq \frac{1}{\Psi_0(y)+1})$, a contradiction. Hence, we have proved $(\forall q^0 \in [0,1])(\exists^{\mathrm{st}} y \in [0,1])(q \approx y)$, which immediately yields $(\forall x^1 \in [0,1])(\exists^{\mathrm{st}} y^1 \in [0,1])(x \approx y)$, as we have $x \approx [x](N)$ for any $x \in [0,1]$ and nonstandard N^0. However, every real has a binary expansion in RCA_0 (See [6]), and B_0^- similarly proves that every (standard) real has a (standard) binary expansion. A real with non-unique binary expansion can be be summed with an infinitesimal to yield a real with a unique binary expansion. Hence, the previous yields that $(\forall \alpha^1 \leq_1 1)(\exists^{\mathrm{st}} \beta^1 \leq_1 1)(\alpha \approx_1 \beta)$, which is just STP. The law of excluded middle as in $(\exists^2) \vee \neg(\exists^2)$ now establishes the reverse implication over $\mathsf{B}_0 + \mathsf{QF\text{-}AC}^{2,1}$.

For the first forward direction, STP implies $(\forall x^1 \in [0,1])(\exists^{\mathrm{st}} y^1 \in [0,1])(x \approx y)$ as in the previous paragraph, and we thus have:

$$(\forall^{\mathrm{st}} \Psi^2)(\forall x^1 \in [0,1])(\exists^{\mathrm{st}} y^1 \in [0,1])(|x-y| < \tfrac{1}{\Psi(y)+1}), \tag{3.9}$$

Applying *Idealisation* to (3.9), we obtain

$$(\forall^{\mathrm{st}} \Psi^2)(\exists^{\mathrm{st}} w^1)(\forall x^1 \in [0,1])(\exists y \in w)(|x-y| < \tfrac{1}{\Psi(y)+1}). \tag{3.10}$$

Dropping the second 'st' in (3.10) and applying $\mathsf{PF\text{-}TP}_\forall$, we obtain HBU.

For the equivalence $\mathsf{HBU}^{\text{st}} \leftrightarrow \mathsf{STP}$, the reverse implication follows from the fact that STP implies (3.10). For the forward implication, note that HBU^{st} implies (3.10) by taking w provided by HBU^{st} and extending this sequence with all $w(i) \pm \Psi(w(i))$ for $i < |w|$. However, (3.10) implies STP by the previous.

Finally, the non-implication follows from [15, Theorem 4.2] as the latter expresses that the special fan functional is not computable in any type two functional. Indeed, STP is equivalent to (3.1) by Theorem 3.2, and apply Theorem 2.6 to $\mathsf{B}_0^- + (\exists^3) + \mathsf{HBU} + \mathsf{QF\text{-}AC} \vdash (3.1)$, to obtain a term t of Gödel's T such that $\mathsf{SCF}(t)$, which is impossible, and we are done. □

Finally, we consider the least-upper-bound princple from [17, §4]. To this end, a formula $\varphi(x^1)$ is called *extensional on* $\mathbb{R}$ if we have

$$(\forall x, y \in \mathbb{R})(x =_{\mathbb{R}} y \to \varphi(x) \leftrightarrow \varphi(y)). \tag{3.11}$$

Note that the same condition is used in RM for defining open sets as in [24, II.5.7].

Principle 3.6 (LUB) *For second-order φ (with any parameters), if $\varphi(x^1)$ is extensional on $\mathbb{R}$ and $\varphi(0) \wedge \neg\varphi(1)$, there is a least $y \in [0,1]$ such that $(\forall z \in (y,1])\neg\varphi(y)$.*

By LUB^{st} we mean LUB with all quantifiers relative to 'st', *including* those pertaining to the parameters in the formula φ^{st}, and all quantifiers in (3.11).

Corollary 3.7 *The system B_0^- proves $\mathsf{LUB}^{\text{st}} \to \mathsf{HBU}^{\text{st}} \to \mathsf{STP}$.*

Proof. The second implication follows by noting that the above proof does not require $\mathsf{QF\text{-}AC}$ or $\mathsf{PF\text{-}TP}_\forall$. The first implication follows from [17, Thm 4.2].

The previous corollary has noteworthy foundational implications: the axiom STP (and the same for LMP from Section 3.3) is what is called a 'purely nonstandard axiom' (See [25, Remark 3.8]). Intuitively speaking, such an axiom does not follow from any true second-order sentence relative to the standard world, i.e. purely nonstandard axioms do not follow from standard axioms. However, STP does follow from the third-order sentence HBU^{st}, as well as from a second-order schema with third-order parameters LUB^{st}. Hence, the notion of "purely nonstandard axiom" is extremely dependent on the exact formal framework.

3.3 Weak compactness and the weak fan functional

Clearly, HBU is a generalisation of WKL from RM. In this section, we list results similar to Theorems 3.4 and 3.5 for generalisations of WWKL. The *weak fan functional* Λ from [15] arises from the axiom WWKL, as follows:

$$(\forall T \leq_1 1)\big[\mu(T) >_{\mathbb{R}} 0 \to (\exists \beta \leq_1 1)(\forall m)(\overline{\beta}m \in T)\big], \tag{WWKL}$$

where '$\mu(T) >_{\mathbb{R}} 0$' is $(\exists k^0)(\forall n^0)\big(\frac{|\{\sigma \in T : |\sigma| = n\}|}{2^n} \geq \frac{1}{k}\big)$. Although WWKL is not part of the Big Five, it sports *some* equivalences ([24, X.1]). The following fragment of *Standard Part* is the nonstandard counterpart of WWKL, as studied in [25]:

$$(\forall T^1 \leq_1 1)\big[\mu(T) \gg 0 \to (\exists^{\text{st}} \beta^1 \leq_1 1)(\forall^{\text{st}} m^0)(\overline{\beta}m \in T)\big], \tag{LMP}$$

where '$\mu(T) \gg 0$' is just the formula $[\mu(T) >_{\mathbb{R}} 0]^{\text{st}}$. Clearly, WWKL and LMP are weakened versions of WKL and STP; the following weaker version of the special fan functional arises from LMP. As for the special one, there is *no unique* weak fan functional, i.e. it is in principle incorrect to refer to 'the' weak fan functional.

Definition 3.8 [Weak fan functional] We define $\mathsf{WCF}(\Lambda)$ for $\Lambda^{(2\to(1\times 1))}$:

$$(\forall k^0, g^2, T^1 \leq_1 1)\big[(\forall \alpha \in \Lambda(g,k)(2))(\overline{\alpha}g(\alpha) \notin T) \to (\exists n \leq \Lambda(g,k)(1))(L_n(T) \leq \tfrac{1}{k})\big].$$

Any Λ satisfying $\mathsf{WCF}(\Lambda)$ is referred to as a *weak fan functional.*

Now, WWKL is equivalent to the following statement: *for every X^1, there is Y^1 which is Martin-Löf random relative to X*, as proved in [1, Theorem 3.1]. This equivalence is proved in RCA_0, and the latter also suffices to e.g. define a *universal* Martin-Löf test $(U_i^X)_{i\in\mathbb{N}}$ (relative to any X^1). The latter has type $0 \to 1$ and represents a universal and effective (relative to X) null set, i.e. a rare event. Intuitively, Y is (Martin-Löf) *random* relative to X, if Y is not in such a rare event. To make this more precise, define '$f^1 \in [\sigma^0]$' as $\overline{f}|\sigma| =_0 \sigma$ for any finite binary sequence and define $\mathsf{MLR}(Y,X)$ as $(\exists i^0)(\forall w^0 \in U_i^X)(Y \notin [w])$.

We can now define restrictions of STP and HBU to Martin-Löf random reals.

$$(\forall^{\text{st}} X^1)(\forall Y^1)(\exists^{\text{st}} Z^1)\big([\mathsf{MLR}(Y,X)]^{\text{st}} \to Z \approx_1 Y\big). \qquad (\mathsf{MLR_{ns}})$$

Let $\mathsf{MLR}(X,Y,i)$ be $\mathsf{MLR}(X,Y)$ without the leading quantifier. Now consider

$$(\forall \Psi^2, k^0, X^1)(\exists w^1)(\forall Y)(\exists Z \in w)(\mathsf{MLR}(Y,X,k) \to Y \in [\overline{Z}\Psi(Z)]). \qquad (\mathsf{HBU_{ml}})$$

Note that $\mathsf{HBU_{ml}}$ expresses that the canonical cover $\cup_{f\in 2^{\mathbb{N}}}[\overline{f}\Psi(f)]$ has a finite sub-cover which covers all reals which are random *and already outside the universal test at level U_k^X* of the universal test. Since $\mu(U_k) \leq \frac{1}{2^k}$, the finite sub-cover need not cover a measure one set in Cantor space. The following theorem is proved in the same way as Theorems 3.4 and 3.5.

Theorem 3.9 *The system* $\mathsf{B}_0 + \mathsf{QF\text{-}AC}^{2,1}$ *proves* $\mathsf{LMP} \leftrightarrow \mathsf{MLR_{ns}} \leftrightarrow \mathsf{HBU_{ml}}$. *Additionally assuming* (FF)*, we also obtain an equivalence to* $(\exists \Lambda)\mathsf{WCF}(\Lambda)$.

Clearly, we may weaken (FF) in the theorem to a functional only implying WWKL.

Finally, the 'st' in the antecedent of LMP (and $\mathsf{MLR_{ns}}$) is essential: in particular, we show that STP (and hence Θ) is *robust* in the sense of RM ([13, p. 495]), but LMP is not. Consider the following variations of LMP and STP.

$$(\forall T \leq_1 1)\big[\mu(T) >_{\mathbb{R}} 0 \to (\exists^{\text{st}} \beta \leq_1 1)(\forall^{\text{st}} m)(\overline{\beta}m \in T)\big], \qquad (\mathsf{LMP}^+)$$

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

where the second one is called 'STP^-'. We have the following theorem.

Theorem 3.10 *In* $\mathsf{B}_0^- + \mathsf{WWKL}$*, we have* $\mathsf{STP} \leftrightarrow \mathsf{LMP}^+ \leftrightarrow \mathsf{STP}^-$.

Proof. For the first equivalence, we only need to prove $\mathsf{STP} \leftarrow \mathsf{STP}^-$, which follows by taking a tree $T \leq_1 1$ as in the antecedent STP, noting that by overspill it has a sequence of nonstandard length, and extending this sequence with $00\dots$ to obtain a tree as in the antecedent of STP^-. Then STP^- yields a standard path in the standard part of the modified tree, which is thus also in the standard part of the original tree. For $\mathsf{STP} \rightarrow \mathsf{LMP}^+$, apply STP to the path claimed to exist by WWKL and note that we obtain LMP^+. For $\mathsf{LMP}^+ \rightarrow \mathsf{STP}$, fix $f^1 \leq_1 1$ and nonstandard N. Define the tree $T \leq_1 1$ which is f until height N, followed by the full binary tree. Then $\mu(T) >_{\mathbb{R}} 0$ and let standard $g^1 \leq_1 1$ be such that $(\forall^{\text{st}} n)(\overline{g}n \in T)$. By definition, $f \approx_1 g$ follows, and we are done. □

References

[1] Jeremy Avigad, Edward T. Dean, and Jason Rute, *Algorithmic randomness, reverse mathematics, and the dominated convergence theorem*, Ann. Pure Appl. Logic **163** (2012), no. 12, 1854–1864.

[2] Benno van den Berg, Eyvind Briseid, and Pavol Safarik, *A functional interpretation for nonstandard arithmetic*, Ann. Pure Appl. Logic **163** (2012), 1962–1994.

[3] Benno van den Berg and Sam Sanders, *Reverse Mathematics and parameter-free Transfer*, To appear in Annals of Pure and Applied Logic (2018). arXiv: `http://arxiv.org/abs/1409.6881`.

[4] Pierre Cousin, *Sur les fonctions de n variables complexes*, Acta Math. **19** (1895), 1–61.

[5] Harvey Friedman, *Some systems of second order arithmetic and their use*, Proceedings of the ICM (Vancouver, B. C., 1974), Vol. 1, 1975, pp. 235–242.

[6] Jeffry L. Hirst, *Representations of reals in reverse mathematics*, Bull. Pol. Acad. Sci. Math. **55** (2007), no. 4, 303–316.

[7] 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.

[8] H. Jerome Keisler, *Nonstandard arithmetic and reverse mathematics*, Bull. Symb. Logic **12** (2006), 100–125.

[9] Ulrich Kohlenbach, *Applied proof theory: proof interpretations and their use in mathematics*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2008.

[10] ———, *Higher order reverse mathematics*, Reverse mathematics 2001, Lect. Notes Log., vol. 21, ASL, 2005, pp. 281–295.

[11] ———, *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.

[12] John Longley and Dag Normann, *Higher-order Computability*, Theory and Applications of Computability, Springer, 2015.

[13] Antonio Montalbán, *Open questions in reverse mathematics*, Bull. Symbolic Logic **17** (2011), no. 3, 431–454.

[14] Edward Nelson, *Internal set theory: a new approach to nonstandard analysis*, Bull. Amer. Math. Soc. **83** (1977), no. 6, 1165–1198.

[15] Dag Normann and Sam Sanders, *Nonstandard Analysis, Computability Theory, and their connections*, arXiv: `https://arxiv.org/abs/1702.06556` (2017).

[16] ———, *Nonstandard Analysis, Computability Theory, and metastability*, In Preparation (2018).

[17] ———, *On the mathematical and foundational significance of the uncountable*, Submitted, https://arxiv.org/abs/1711.08939 (2017).
[18] Salvatore Pincherle, *Sopra alcuni sviluppi in serie per funzioni analitiche (1882)*, Opere Scelte, I, Roma (1954), 64–91.
[19] Sam Sanders, *The Gandy-Hyland functional and a computational aspect of Nonstandard Analysis*, *Computability*, arXiv: http://arxiv.org/abs/1502.03622 (2018).
[20] ———, *Formalism16*, Synthese, S.I.: Foundations of Mathematics (2017), 1-42.
[21] ———, *Metastability and higher-order computability*, Proceedings of LFCS18, Lecture Notes in Computer Science, Springer (2018).
[22] ———, *Some nonstandard equivalences in Reverse Mathematics*, Proceedings of CiE18, Lecture Notes in Computer Science, Springer (2018).
[23] ———, *To be or not to be constructive*, Special issue of *Indagationes Mathematicae*; https://arxiv.org/abs/1704.00462 (2017), pp. 68.
[24] Stephen G. Simpson, *Subsystems of second order arithmetic*, 2nd ed., Perspectives in Logic, CUP, 2009.
[25] Stephen G. Simpson and Keita Yokoyama, *A nonstandard counterpart of* WWKL, Notre Dame J. Form. Log. **52** (2011), no. 3, 229–243.