THE STRENGTH OF COMPACTNESS IN COMPUTABILITY THEORY AND NONSTANDARD ANALYSIS

DAG NORMANN AND SAM SANDERS

ABSTRACT. Compactness is one of the core notions of analysis: it connects local properties to global ones and makes limits well-behaved. We study the computational properties of the compactness of Cantor space $2^{\mathbb{N}}$ *for uncountable covers*. The most basic question is: *how hard is it to compute a finite sub-cover from such a cover of* $2^{\mathbb{N}}$? Another natural question is: *how hard is it to compute a sequence that covers* $2^{\mathbb{N}}$ *minus a measure zero set from such a cover*? The *special* and *weak* fan functionals respectively compute such finite sub-covers and sequences. In this paper, we establish the connection between these new fan functionals on one hand, and various well-known comprehension axioms on the other hand, including arithmetical comprehension, transfinite recursion, and the Suslin functional. In the spirit of *Reverse Mathematics*, we also analyse the logical strength of compactness in *Nonstandard Analysis*. Perhaps surprisingly, the results in the latter mirror (often perfectly) the computational properties of the special and weak fan functionals. In particular, we show that compactness (nonstandard or otherwise) readily brings us to the outer edges of Reverse Mathematics (namely Π_2^1-CA_0), and even into Schweber's higher-order framework (namely Σ_1^2-separation).

1. INTRODUCTION

The importance of (open-cover) *compactness* can hardly be overstated, as it allows one to treat uncountable sets like Cantor space as 'almost finite' by connecting local properties to global ones. A famous example is *Heine's theorem*, i.e. the *local* property of continuity implies the *global* property of *uniform* continuity on the unit interval. In general, Tao writes:

> Compactness is a powerful property of spaces, and is used in many ways in many different areas of mathematics. One is via appeal to local-to-global principles; one establishes local control on some function or other quantity, and then uses compactness to boost the local control to global control. [57, p. 168]

Compactness already has a long history: the *Cousin lemma* ([6, p. 22]) on the open-cover compactness of subsets of $\mathbb{R}^2$, dates back[1] 135 years. Despite its basic

[1]The collected works of Pincherle contain a footnote (see [38, p. 67]) which states that the associated *Teorema* from 1882 corresponds to the Heine-Borel theorem. This claim is repeated in [37]. Moreover, Weierstrass proves the Heine-Borel theorem (without explicitly formulating it) in 1880 in [61, p. 204]. A detailed motivation for these claims may be found in [22, p. 96-97].

nature, its central role in analysis, and a long history, little is known about the logical and computational properties of compactness. The main aim of this paper is to study these computational properties. In particular, we are interested in the following most basic and natural question (and its variations):

Given an uncountable cover of $2^{\mathbb{N}}$, how hard is it to compute a finite sub-cover?

To answer this question, we continue the project initiated in [33], namely we study the computational properties of *special fan functionals* (and their variations). The latter compute the aforementioned finite sub-covers, as detailed in (T.1) below. In the spirit of *Reverse Mathematics*, we also analyse the logical strength of compactness in *Nonstandard Analysis* as in (T.2) below. As it happens, the results in Nonstandard Analysis mirror (often perfectly) the results in Computability Theory. We assume basic familiarity with the aforementioned fields, in particular the program Reverse Mathematics founded by Friedman (RM hereafter; see [53, 54, 56] or [33, §2.2]). In Section 2, we provide an overview of the results in [33], and a list of the questions to be answered, all pertaining to the following two topics. Many questions left open in, or raised by, [33] are in fact answered in this paper. We refer to [33, 34] for an introduction and overview to the project this paper is part of. In this paper, we explore the following topics:

Topic (T.1): We study two new classes of functionals, namely the *special fan functionals*, an instance of which is denoted Θ, and the (computationally weaker) *weak fan functionals*, an instance of which is denoted Λ. Intuitively speaking, any Θ computes a finite sub-cover from an uncountable cover of Cantor space, while any Λ provides such a cover 'in the limit'. These functionals are quite natural mathematical objects: The special fan functionals emerge *naturally and directly* from Tao's *metastability* ([50]) while the existence of Θ is equivalent to *Cousin's lemma* ([34, §3.3]), and to many basic properties of the *gauge integral*; the latter in turn provides a unique/direct[2] formalisation of Feyman's path integral ([24]). From the perspective of higher-order computability theory, these new fan functionals are interesting as they fall *outside* the well-studied classes, like e.g. the continuous functionals or the so-called normal functionals. In this paper, we establish the connection between Λ and Θ on one hand, and arithmetical comprehension, transfinite recursion, and the Suslin functional on the other hand. The new fan functionals will be seen to exhibit rather surprising behaviour.

Topic (T.2): We study the *nonstandard counterparts* of the 'Big Five' systems WKL_0, ACA_0, and $\Pi^1_1\text{-}\mathsf{CA}_0$ of RM. These counterparts are respectively: the nonstandard compactness of Cantor space STP, the *Transfer* axiom limited to Π^0_1-formulas $\Pi^0_1\text{-}\mathsf{TRANS}$, and the *Transfer* axiom limited to Π^1_1-formulas $\Pi^1_1\text{-}\mathsf{TRANS}$. While the original Big Five systems are linearly ordered as follows

$$\Pi^1_1\text{-}\mathsf{CA}_0 \rightarrow \mathsf{ATR}_0 \rightarrow \mathsf{ACA}_0 \rightarrow \mathsf{WKL}_0 \rightarrow \mathsf{RCA}_0,$$

the non-implications $\Pi^0_1\text{-}\mathsf{TRANS} \nrightarrow \mathsf{STP} \nleftarrow \Pi^1_1\text{-}\mathsf{TRANS}$ hold for the respective nonstandard counterparts, as proved in [33]. In this paper, we study the strength of

---

[2]There are a number of different approaches to the formalisation of Feynman's path integral. However, if one requires the formalisation to be close to Feynman's original formulation, then the gauge integral is the only approach (see [34, §3.3] for a discussion). Another argument in favour of the gauge integral is that this formalism gives rise to so-called physical solutions, i.e. in line with the observations from physics (see [28–30, 39]), in particular the absence of 'imaginary time'.

Π_1^1-TRANS + STP which (indirectly) yields results about the strength of the combination of Θ and the Suslin functional. We study Schweber's third-order framework ([51, 52]) via Nonstandard Analysis and obtain some results involving *compactness of function spaces*. While interesting in its own right, the aforementioned compactness is essential to the *gauge integral* over function spaces, which in turn formalises the *Feynman path integral*.

As it turns out, topics (T.1) and (T.2) are intimately connected: (non-) computability results in (T.1) are obtained *directly* from (non-) implications in (T.2), and vice versa. In fact, Θ first arose from nonstandard compactness as in STP when studying the computational content of Nonstandard Analysis ([44]), while instances of the axiom *Transfer* give rise to (well-known) comprehension and choice functionals. As it happens, the connection between Θ and metastability was first proved *via Nonstandard Analysis* ([50]). It should be noted that our definition of these new fan functionals, to be found in Section 2.1, is *different* from the (original) definition used in e.g. [44]. The definitions are equivalent as shown in Section 2.6.

We now sketch the main results of this paper as follows. A detailed discussion may be found in Section 2.5. Feferman's μ^2 is introduced in Section 2.3 and constitutes a form of arithmetical comprehension.

(i) The Suslin functional is not computable from $\Theta + \mu^2$ (Section 3.2). The combination $\Theta + \mu^2$ (directly) computes a realiser for ATR_0 (Section 3.3).

(ii) The combination Π_1^1-TRANS + STP exists at the level of Π_2^1-CA_0 (Section 4.1). This result yields results not involving Nonstandard Analysis.

(iii) We identify a weak fan functional Λ_1 and show in Section 3.4 that $\Lambda_1 + \mu^2$ computes the same objects as μ^2. This shows that we cannot in general compute a special fan functional from a weak one.

(iv) We show that some of our results, Theorem 2.19 in particular, generalise to Schweber's third-order arithmetic [51, 52] (Section 4.2).

Finally, this paper connects Computability Theory and Nonstandard Analysis. The first author contributed most results in the former, while the second author did so for the latter. However, many questions were answered by translating them from one field to the other, solving them, and translating everything back, i.e. both authors contributed somehow to most of the paper. As suggested by the above, this paper is part of a series of papers by the authors, as follows. In our first two papers ([33] and this paper) we link Nonstandard Analysis and higher order Computability Theory, while in the other three ([34–36]) we focus on the logical and computational content of classical theorems in mathematical analysis.

2. Previous work and open questions

We introduce the weak and special fan functionals and discuss their connection to nonstandard compactness. We discuss the associated results in Computability Theory and Nonstandard Analysis from [33] and list the open questions to be answered below. We first make our notion of 'computability' precise as follows.

(I) We adopt ZFC set theory as the official metatheory for all results, unless explicitly stated otherwise.

(II) We adopt Kleene's notion of *higher-order computation* as given by his nine clauses S1-S9 (see [20, 43]) as our official notion of 'computable'.

In Section 3, we provide the basic definitions of Computability Theory needed for (II), but do assume some familiarity with Computability Theory as a whole. We refer to [33, §2] or [48] for an introduction to Nelson's system IST and the fragments P and P_0 which are conservative extensions of Peano arithmetic and RCA_0. For completeness, the systems P and P_0 can be found in Appendix A.

Finally, to improve readability, we often omit types if they can be gleaned from context; we sometimes make use of set theoretical notation. For instance, '$\alpha^1 \leq 1$' expresses that α is a binary sequence, but could also be written $\alpha \leq 1$ or $\alpha \in 2^{\mathbb{N}}$ or $\alpha \in C$. Details regarding the former notation may be found in Notation A.2.

2.1. **The special and weak fan functionals.** First of all, we define two new classes of functionals. The *special fan functionals* intuitively output a finite sub-cover on input an uncountable cover of $2^{\mathbb{N}}$. The (computationally weaker) *weak fan functionals* take an additional input $k \in \mathbb{N}$ and output a finite sub-cover for a subset of $2^{\mathbb{N}}$ of measure at least $1 - \frac{1}{2^k}$. We usually simplify the type of these fan functionals to '3'. We reserve the symbols Θ and Λ to denote instances of the special and weak fan functionals. It goes without saying these functionals are not unique: one can always add extra binary sequences to the finite sub-cover.

We now introduce the class of special fan functionals. We write '$f \in [\sigma]$' for $\overline{f}|\sigma| =_{0^*} \sigma$, where τ^* is the type of finite sequences of type τ objects. For $w^{\tau^*} = \langle t_0, \ldots, t_k \rangle$, we write $|w| = k + 1$ and $w(i) = t_i$ for $i < |w|$. These 'finite sequence' notations are discussed in detail in Notation A.1.

Definition 2.1. [Special fan functionals] $\mathsf{SFF}(\Theta)$ is as follows for $\Theta^{2 \to 1^*}$:

$$(\forall G^2)(\forall f^1 \leq 1)(\exists g \in \Theta(G))(f \in [\overline{g}G(g)]), \tag{2.1}$$

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

Intuitively, any functional G^2 gives rise to a 'canonical cover' $\cup_{f \in 2^{\mathbb{N}}}[\overline{f}G(f)]$ of Cantor space, and $\Theta(G)$ is a finite sub-cover thereof, i.e. $\cup_{g \in \Theta(G)}[\overline{g}G(g)]$ also covers $2^{\mathbb{N}}$. Note that Cousin ([6]) and Lindelöf ([19]) make use of such canonical covers (for $\mathbb{R}^n$) rather than the modern/general notion of cover. In light of (2.1), special fan functionals may be called 'realisers for the Heine-Borel theorem or Cousin lemma for C'. As it happens, Θ actually arises from the *nonstandard compactness of C* as in *Robinson's theorem* ([13, p. 42]), as discussed in Sections 2.2 and 2.6.

Secondly, we introduce the class of *weak fan functionals* Λ, which are strictly weaker than Θ in general. As will become clear below, Λ is not just 'more of the same' but occupies an important place relative to Θ. Where $\Theta(G)$ provides a finite sub-cover of C, $\Lambda(G, k)$ only yields a finite sub-cover of a subset of C with measure at least $1 - \frac{1}{2^k}$, i.e. we have the following:

$$\mathbf{m}(\{f \in C : (\exists g \in \Lambda(G, k))(f \in [\overline{g}G(g)])\}) \geq 1 - \frac{1}{2^k}, \tag{2.2}$$

where $\mathbf{m}$ is the usual coin-toss measure on $2^{\mathbb{N}}$. It is straightforward, but cumbersome, to formally express (2.2) in our formal language.

Definition 2.2. [Weak fan functionals] $\mathsf{WFF}(\Lambda)$ is as follows for $\Lambda^{(2 \times 0) \to 1^*}$:

$$(\forall G^2, k^0)\big[\mathbf{m}(\{f \in C : (\exists g \in \Lambda(G, k))(f \in [\overline{g}G(g)])\}) \geq 1 - \frac{1}{2^k}\big]. \tag{2.3}$$

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

Weak fan functionals are not realisers of theorems from the literature, but these functionals do capture the core complexity of several theorems concerning measure-theoretic approximations, like the Vitali Covering Theorem ([60]). This is investigated further in [36]. As it happens, weak fan functionals also arise from *nonstandard compactness*, as discussed in Sections 2.2 and 2.6.

Finally, Θ appears similar in name and behaviour to Tait's 'classical' fan functional (esp. on the continuous functionals). However, Θ and Λ behave quite differently in that they cannot be computed by *any* type two functional (see Section 2.3).

2.2. **Nonstandard compactness and related notions.** In this section, we introduce some axioms of Nonstandard Analysis. We will observe that the special and weak fan functionals emerge from the *nonstandard compactness of Cantor space*.

First of all, we mention the crucial theorem which connects P and Peano arithmetic. Definitions may be found in [4], [33, §2], [48, Appendix], or Appendix A

Theorem 2.3 (Term extraction). *If Δ_{int} is a collection of internal formulas and ψ is internal, and*

$$\mathsf{P} + \Delta_{\mathsf{int}} \vdash (\forall^{\mathrm{st}}\underline{x})(\exists^{\mathrm{st}}\underline{y})\psi(\underline{x}, \underline{y}, \underline{a}), \tag{2.4}$$

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

$$\mathsf{E\text{-}PA}^{\omega*} + \Delta_{\mathsf{int}} \vdash (\forall\underline{x})(\exists\underline{y} \in t(\underline{x}))\psi(\underline{x}, \underline{y}, \underline{a}). \tag{2.5}$$

Proof. See [44, §2] or [48, Appendix]. The route from (2.4) to (2.5) involves a functional interpretation called 'S_{st}', introduced in [4]. $\qquad\square$

The system $\mathsf{RCA}_0^\omega \equiv \mathsf{E\text{-}PRA}^\omega + \mathsf{QF\text{-}AC}^{1,0}$ is Kohlenbach's *base theory of higher-order Reverse Mathematics* as introduced in [17, §2]. We permit ourselves a slight abuse of notation by also referring to the system $\mathsf{E\text{-}PRA}^{\omega*} + \mathsf{QF\text{-}AC}^{1,0}$ as RCA_0^ω.

Corollary 2.4. *The previous theorem and corollary go through for P and $\mathsf{E\text{-}PA}^{\omega*}$ replaced by $\mathsf{P}_0 \equiv \mathsf{E\text{-}PRA}^{\omega*} + \mathcal{T}_{\mathrm{st}}^* + \mathsf{HAC}_{\mathsf{int}} + \mathsf{I} + \mathsf{QF\text{-}AC}^{1,0}$ and RCA_0^ω.*

From now on, the notion 'normal form' refers to a formula as in (2.4), i.e. of the form $(\forall^{\mathrm{st}}x)(\exists^{\mathrm{st}}y)\varphi(x, y)$ for φ internal. We now provide a general template how term extraction is used below, as this will shorten a number of proofs.

Remark 2.5 (Using term extraction). First of all, term extraction as in Theorem 2.3 is restricted to normal forms. We now show that normals forms are 'closed under implication', as follows. Let φ, ψ be internal and consider the following implication between normal forms:

$$(\forall^{\mathrm{st}}x)(\exists^{\mathrm{st}}y)\varphi(x, y) \to (\forall^{\mathrm{st}}z)(\exists^{\mathrm{st}}w)\psi(z, w). \tag{2.6}$$

Since standard functionals have standard output for standard input, (2.6) implies

$$(\forall^{\mathrm{st}}\zeta)\big[(\forall^{\mathrm{st}}x)\varphi(x, \zeta(x)) \to (\forall^{\mathrm{st}}z)(\exists^{\mathrm{st}}w)\psi(z, w)\big]. \tag{2.7}$$

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

$$(\forall^{\mathrm{st}}\zeta, z)(\exists^{\mathrm{st}}w, x)\big[\varphi(x, \zeta(x)) \to \psi(z, w)\big], \tag{2.8}$$

as the formula in square brackets is internal. Now, (2.8) is equivalent to (2.7), but one usually weakens the latter as follows:

$$(\forall^{\mathrm{st}}\zeta, z)(\exists^{\mathrm{st}}w)\big[(\forall x)\varphi(x, \zeta(x)) \to \psi(z, w)\big], \tag{2.9}$$

as (2.9) is closer to the usual mathematical definitions.

Secondly, assuming (2.6) is provable in P, so is (2.9) and we obtain a term t with

$$(\forall \zeta, z)(\exists w \in t(\zeta, z))\big[(\forall x)\varphi(x, \zeta(x)) \to \psi(z, w)\big] \tag{2.10}$$

being provable in $\mathsf{E\text{-}PA}^{\omega*}$. We now omit the term t and bring all quantifiers inside again, yielding that $\mathsf{E\text{-}PA}^{\omega*}$ proves:

$$(\exists \zeta)(\forall x)\varphi(x, \zeta(x)) \to (\forall z)(\exists w)\psi(z, w). \tag{2.11}$$

Finally, we shall often shorten the below proofs by just providing normal forms and jumping straight from (2.6) to (2.11) whenever possible.

Secondly, P does not involve Nelson's axiom *Transfer*, as 'small' fragments are already quite strong. Indeed, *Transfer* restricted to Π_1^0-formulas as follows

$$(\forall^{\mathrm{st}} f^1)\big[(\forall^{\mathrm{st}} n)f(n) \neq 0 \to (\forall m)f(m) \neq 0\big] \tag{Π_1^0-TRANS}$$

is the nonstandard counterpart of arithmetical[3] comprehension as in ACA_0. Furthermore, the fragment[4] of *Transfer* for Π_1^1-formulas as follows

$$(\forall^{\mathrm{st}} f^1)\big[(\exists g^1)(\forall n^0)(f(\bar{g}n) = 0) \to (\exists^{\mathrm{st}} g^1)(\forall n^0)(f(\bar{g}n) = 0)\big] \tag{Π_1^1-TRANS}$$

is the nonstandard counterpart of Π_1^1-CA_0. It is an interesting exercise to show that if the antecedent of (2.6) is Π_1^0-TRANS (resp. Π_1^1-TRANS), the antecedent of (2.11) is (μ^2) (resp. (μ_1)), to be introduced in Section 2.3.

The following fragment of *Standard Part* is the nonstandard counterpart of weak König's lemma ([14]):

$$(\forall \alpha^1 \leq 1)(\exists^{\mathrm{st}} \beta^1 \leq 1)(\alpha \approx_1 \beta), \tag{STP}$$

where $\alpha \approx_1 \beta$ is $(\forall^{\mathrm{st}} n)(\alpha(n) =_0 \beta(n))$. Note that STP expresses the *nonstandard compactness of $2^{\mathbb{N}}$* as in *Robinson's theorem* ([13, p. 42]), The following fragment of *Standard Part* is the nonstandard counterpart of weak weak König's lemma ([55]). We reserve the variable 'T^1' for trees and '$T^1 \leq 1$' means that T is a binary tree.

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

where '$\mu(T) \gg 0$' is just the formula $(\exists^{\mathrm{st}} k^0)(\forall^{\mathrm{st}} n^0)\left(\frac{\{\sigma \in T : |\sigma| = n\}}{2^n} \geq_0 \frac{1}{2^k}\right)$.

Note that there is no deep philosophical meaning to be found in the words 'non-standard counterpart': this is just what the principles STP, LMP, Π_1^0-TRANS, and Π_1^1-TRANS are called in the literature ([14, 46, 55]). The following theorems from [33] provide normal forms for STP and LMP and establish the latter's relationships with the special and weak fan functionals. In particular, the latter emerge from STP and LMP when applying Theorem 2.3. Recall the 'finite sequence' notations from Notation A.1.

Theorem 2.6. *In P_0, STP is equivalent to the following:*

$$(\forall^{\mathrm{st}} g^2)(\exists^{\mathrm{st}} w^{1^*} \leq 1, k^0)(\forall T^1 \leq 1)\big[(\forall \alpha^1 \in w)(\bar{\alpha}g(\alpha) \notin T) \tag{2.12}$$
$$\to (\forall \beta \leq 1)(\exists i \leq k)(\bar{\beta}i \notin T)\big],$$

---

[3]Similar to how one 'bootstraps' Π_1^0-comprehension to the latter, the system $\mathsf{P}_0 + \Pi_1^0$-TRANS proves $\varphi \leftrightarrow \varphi^{\mathrm{st}}$ for any internal arithmetical formula (only involving standard parameters).

[4]The 'bootstrapping' trick for Π_1^0-TRANS does not work for Π_1^1-TRANS (or Π_1^1-CA_0) as the latter is restricted to type one objects (like g^1 in Π_1^1-TRANS) occurring as 'call by value'.

and is equivalent to $(\forall^{st}G^2)(\exists^{st}w^{1^*})(\forall f^1 \leq 1)(\exists g \in w)(f \in [\overline{g}G(g)])$, *and to:*

$$(\forall T^1 \leq 1)\big[(\forall^{st}n^0)(\exists\beta^{0^*})(|\beta| = n \wedge \beta \in T) \to (\exists^{st}\alpha^1 \leq 1)(\forall^{st}n^0)(\overline{\alpha}n \in T)\big]. \quad (2.13)$$

Furthermore, P_0 *proves* $(\exists^{st}\Theta)\mathsf{SFF}(\Theta) \to \mathsf{STP}$.

Proof. All results are established in [33], except the following equivalence:

$$\mathsf{STP} \leftrightarrow (\forall^{st}G^2)(\exists^{st}w^{1^*})(\forall f^1 \leq 1)(\exists g \in w)(f \in [\overline{g}G(g)]). \quad (2.14)$$

To establish (2.14), use $\mathsf{HAC}_{\mathsf{int}}$ to establish that $(\exists^{st}g \leq 1)(\forall^{st}k^0)(\overline{f}k =_0 \overline{g}k)$ is equivalent to $(\forall^{st}G^2)(\exists^{st}g \leq 1)(\overline{f}G(g) =_{0^*} \overline{g}G(g))$ (by considering the negations of the latter two formulas). Now prepend '$(\forall f^1 \leq 1)$' to the latter formula and use *Idealisation* to pull the '$(\exists^{st}g \leq 1)$' to the front as in (2.14). $\qquad\square$

By (2.13) in the theorem, STP is just WKL^{st} with the leading 'st' dropped; this observation explains why STP deserves the monicker 'nonstandard counterpart of WKL'. The following theorem follows in the same way.

Theorem 2.7. *In* P_0, *the principle* LMP *is equivalent to:*

$$(\forall^{st}g^2, k^0)(\exists^{st}w^{1^*} \leq 1, n^0)(\forall T \leq 1)\big[(\forall\alpha \in w)(\overline{\alpha}g(\alpha) \notin T) \to \frac{|\{\sigma \in T : |\sigma| = n\}|}{2^n} \leq \frac{1}{2^k}\big].$$

Furthermore, P_0 *proves* $(\exists^{st}\Lambda)\mathsf{WFF}(\Lambda) \to \mathsf{LMP}$.

Despite STP and LMP being the nonstandard counterparts of WKL and WWKL, the former behaves *quite* differently from the latter (and (2.15)) in that the former does not follow from $\Pi_1^0\text{-}\mathsf{TRANS}$ or $\Pi_1^1\text{-}\mathsf{TRANS}$, i.e. the nonstandard counterparts of ACA_0 and $\Pi_1^1\text{-}\mathsf{CA}_0$, as discussed in Section 2.4.

Finally, we discuss the exact connection between our systems of Nonstandard Analysis and Computability theory provided by Theorem 2.3. The crucial point here is that in the syntactic theory of Nonstandard Analysis, the usual quantifiers $\exists$ and $\forall$ play the role of 'uniform quantifiers' (see [5]) which are *ignored* by the functional interpretation S_{st} used in the proof of Theorem 2.3, while the standard quantifiers $\exists^{st}$ and $\forall^{st}$ are given computational meaning. Indeed, the functional interpretation S_{st} applied to the proof of (2.4) yields a term $t(\underline{x})$ in which the $(\forall^{st}\underline{x})$ quantifier in (2.4) describes the input variables, while the $(\exists^{st}\underline{y})$ quantifier describes the output variables. This gives each of the nonstandard axioms a clear computational meaning entirely independent of Nonstandard Analysis per se, which may be of comfort to some who find Nonstandard Analysis alien. Those interested in this kind of development should consult [48].

2.3. **Known results in Computability Theory.** A substantial number of results regarding the special and weak fan functionals were obtained in [33, 34, 50], some of which we list in this section as they are needed below or give rise to open questions. We recall an oft-made observation regarding WWKL_0 and the 'Big Five' of RM, namely that these six systems satisfy the strict implications:

$$\Pi_1^1\text{-}\mathsf{CA}_0 \to \mathsf{ATR}_0 \to \mathsf{ACA}_0 \to \mathsf{WKL}_0 \to \mathsf{WWKL}_0 \to \mathsf{RCA}_0. \quad (2.15)$$

We mention (2.15) as our results show that the situation is quite different in a higher-order or nonstandard setting. More results of this nature are in [34–36].

First of all, it turns out that the fan functionals Θ and Λ are hard to compute.

Theorem 2.8. *Let* φ^2 *be a type two functional. There is no functional* Θ^3 *as in* $\mathsf{SFF}(\Theta)$ *and no functional* Λ^3 *as in* $\mathsf{WFF}(\Lambda)$ *computable in* φ.

Proof. Immediate from [33, Cor. 3.8 and Theorem 3.14].   □

We now list some well-known type two functionals which will be needed below. *Feferman's search operator* as in (μ^2) (see e.g. [1, §8]) is equivalent to $(\exists^2)$ over Kohlenbach's system RCA_0^ω by [18, §3]:

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

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

Furthermore, $\mathsf{ACA}_0^\omega \equiv \mathsf{RCA}_0^\omega + (\mu^2)$ is a Π_2^1-conservative extension of ACA_0 ([42, Theorem 2.2]). The *Suslin functional* (S^2) and the related (μ_1) (see [1, §8.4.1], [17, §1], and [42, §3]) give rise to Π_1^1-CA_0:

$$(\exists\mu_1^{1\to1})(\forall f^1)\big[(\exists g^1)(\forall x^0)(f(\overline{g}x)=0)\to (\forall x^0)(f(\overline{\mu_1(f)}x)=0)\big]. \qquad (\mu_1)$$

$$(\exists S^2)(\forall f^1)\big[(\exists g^1)(\forall n^0)(f(\overline{g}n)=0)\leftrightarrow S(f)=0\big]. \qquad (S^2)$$

In fact, Π_1^1-$\mathsf{CA}_0^\omega \equiv \mathsf{RCA}_0^\omega + (\mu_1)$ is a Π_3^1-conservative extension of Π_1^1-CA_0 ([42, Theorem 2.2]). We let $\mathsf{SU}(S)$ and $\mathsf{MUO}(\mu_1)$ be (S^3) and (μ_1) without the leading existential quantifiers. Similarly, we introduce Π_k^1-$\mathsf{CA}_0^\omega \equiv \mathsf{RCA}_0^\omega + (S_k^2)$, where (S_k^2) states the existence of a type two function S_k^2 which decides Π_k^1-formulas; note that S_1 is the Suslin functional. The higher-order version of second-order arithmetic Z_2 is $\mathsf{Z}_2^\Omega \equiv \mathsf{RCA}_0^\omega + (\exists^3)$, where the latter is

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

Note that Z_2^Ω and Z_2 prove the same sentences by [12, §2]. We reserve '$\exists^3$' for the unique functional ξ^3 from $(\exists^3)$. We do the same for other functionals, like $\mu^2, \mu_1, S^2, \dots$ introduced above.

Theorem 2.9. *A functional Θ^3 as in $\mathsf{SFF}(\Theta)$ can be computed from $\exists^3$.*

Proof. Immediate from [33, Theorem 3.9].   □

By the following theorem, the exotic properties of Θ are not due to its high type. As discussed in [34], HBU is essentially *Cousin's lemma*, dating as far back as 1882.

Theorem 2.10. $\mathsf{ACA}_0^\omega + \mathsf{QF\text{-}AC}^{2,1}$ *proves* $(\exists\Theta)\mathsf{SFF}(\Theta) \leftrightarrow \mathsf{HBU}$; *the latter is*

$$(\forall\Psi^2 : \mathbb{R}\to\mathbb{R}^+)(\exists w^{1^*})(\forall x^1\in[0,1])(\exists y\in w)(x\in I_y^\Psi), \qquad (\mathsf{HBU})$$

where I_y^Ψ is $(y-\Psi(y), y+\Psi(y))$. No system Π_k^1-CA_0^ω proves HBU.

Proof. Immediate from [34, Theorems 3.1 and 3.3].   □

A similar result can be obtained for Λ: the existence of the latter is equivalent to the fact that a finite sub-cover exists for any open cover of the Martin-Löf random reals in Cantor space *minus* some U_k, where the latter is the k-th set in the universal Martin-Löf test (see [45]). This result originates from the RM of WWKL as in [2].

Theorem 2.10 already deals a significant blow to the elegant picture in (2.15), but HBU can even collapse part of the latter linear order, namely as in Theorem 2.11. Now, ATR_0 is ACA_0 plus *arithmetical transfinite recursion* as follows:

$$(\forall X^1)\big[\mathsf{WO}(X)\to(\exists Y^1)H_\theta(X,Y)\big], \qquad (\mathsf{ATR}_\theta)$$

for any arithmetical θ. Here, $\mathsf{WO}(X)$ expresses that X is a countable well-ordering and $H_\theta(X,Y)$ expresses that Y is the result from iterating θ along X. Details and

definitions may be found in [54, V.2]. For Theorem 2.11, we need the following 'trivially uniform' version of ATR_0:

$$(\exists \Phi^{1\to 1})(\forall X^1, f^1)\big[\mathsf{WO}(X) \to H_f(X, \Phi(X, f))\big], \qquad (\mathsf{UATR})$$

where $H_f(X, Y)$ is just $H_\theta(X, Y)$ with $\theta(n, Z)$ defined as $(\exists m^0)(f(n, m, \overline{Z}m) = 0)$. Note that the base theory in the following theorem is conservative over WKL_0.

Theorem 2.11. *The system* $\mathsf{RCA}_0^\omega + \mathsf{HBU} + \mathsf{QF\text{-}AC}^{2,1}$ *proves* $(\mu^2) \leftrightarrow \mathsf{UATR}$.

Proof. Immediate from [33, Cor. 6.7] and [34, Theorem 3.3]. $\qquad\qquad\square$

The previous theorem is based on an effective result where Φ as in UATR_0 is defined from Θ and μ^2 via a term of Gödel's T. This effective result in turn derives from Theorem 2.19, i.e. via term extraction applied to Nonstandard Analysis.

Theorem 2.12. $\mathsf{RCA}_0^\omega + (\exists \Theta)\mathsf{SFF}(\Theta)$ *is a conservative extension of* $\mathsf{RCA}_0^2 + \mathsf{WKL}$.

Proof. Immediate from [33, Cor. 3.5]. $\qquad\qquad\square$

Combining Theorems 2.10 and 2.11, it would seem that Θ produces non-hyperarithmetical outputs, which turns out to be correct. By contrast, there are weak instances of Λ which are 'closed on the hyperarithmetical'.

Theorem 2.13. *For any* Θ *such that* $\mathsf{SFF}(\Theta)$, *there is hyperarithmetical* G^2 *such that* $\Theta(G)$ *is not hyperarithmetical.*

Proof. Immediate from [33, Theorem 5.1]. $\qquad\qquad\square$

Theorem 2.14. *There is a* Λ_0 *such that* $\mathsf{WFF}(\Lambda_0)$ *and such that for any total, hyperarithmetical* G^2, $\Lambda_0(G, k)$ *is a finite list of hyperarithmetical functions.*

Proof. Immediate from [33, Cor. 5.14]. $\qquad\qquad\square$

As noted in [33, Appendix], and proved in full in Section 3.4, one can define a weak fan functional that is very weak as follows:

Theorem 2.15. *There exists a functional* Λ_1 *satisfying* $\mathsf{WFF}(\Lambda_1)$ *such that all functions computable in* Λ_1 *and* $\exists^2$ *are hyperarithmetical.*

Proof. The proof is given in Section 3.4. See Theorem 3.31. $\qquad\qquad\square$

Corollary 2.16. *There exists a functional* Λ_1 *satisfying* $\mathsf{WFF}(\Lambda_1)$ *such that no* Θ *satisfying* $\mathsf{SFF}(\Theta)$ *is computable in* Λ_1 *and* $\exists^2$.

Proof. Theorems 2.13 and 2.15 immediately yield the corollary. $\qquad\qquad\square$

Finally, Theorem 2.12 is proved using the ECF-*translation*, which will be needed below. We therefore discuss the proof of the former theorem in some detail.

Remark 2.17 (ECF-translation and Θ)**.** As discussed in [17, §3], one can modify the proofs in [58, §2.6] to establish that $\mathsf{RCA}_0^\omega + (\exists \Omega^3)\mathsf{MUC}(\Omega)$ is conservative over $\mathsf{RCA}_0^2 + \mathsf{WKL}$, where Ω^3 is called the *intuitionistic fan functional* as follows:

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

In the latter reference, the so-called ECF-interpretation is defined which, intuitively speaking, replaces all higher-order functionals (of type two or higher) by type one

codes (in the sense of Reverse Mathematics) which *represent* (automatically continuous) higher-type functionals. The ECF-interpretation has the following convenient property (discussed in [17, §3]) for any formula in the language of finite types:

$$\text{If } \mathsf{RCA}_0^\omega \vdash A, \text{ then } \mathsf{RCA}_0^2 \vdash [A]_{\mathsf{ECF}}. \tag{2.16}$$

Now, the ECF-interpretation of $(\exists\Omega^3)\mathsf{MUC}(\Omega)$ expresses that there is a code α^1 which yields a modulus of uniform continuity on Cantor space on input a code β^1 representing an (automatically continuous) type two functional. As follows from the discussion in [20, p. 459], we have $[(\exists\Omega^3)\mathsf{MUC}(\Omega)]_{\mathsf{ECF}} \leftrightarrow \mathsf{WKL}$. Alternatively, one can explicitly define the aforementioned code α^1 and show that it has the required properties using (the contraposition of) WKL, as done in [58, 2.6.6] and [32, p. 101].

Theorem 2.12 can now be obtained in at least two ways: First of all, one considers $(\exists\Omega)\mathsf{MUC}(\Omega) \to (\exists\Theta)\mathsf{SFF}(\Theta) \to \mathsf{WKL}$ (provable in RCA_0^ω), which follows from the results in [33, §3] or [44, §3], and applying the ECF-translation and the above results yields $\mathsf{WKL} \to [(\exists\Theta)\mathsf{SFF}(\Theta)]_{\mathsf{ECF}} \to \mathsf{WKL}$. Secondly, one can also explicitly define the code for Θ required for $[(\exists\Theta)\mathsf{SFF}(\Theta)]_{\mathsf{ECF}}$ in terms of the aforementioned code α^1, as the classical fan functional trivially computes $\Theta(G)$ in case G^2 is continuous on Cantor space. This finishes the proof of Theorem 2.12.

2.4. **Known results in Nonstandard Analysis.** A substantial number of results regarding nonstandard compactness were obtained in [33], some of which we list in this section as they are needed below or give rise to open questions.

First of all, although the Big Five and WWKL_0 are linearly ordered as in (2.15), the nonstandard counterparts behave quite differently.

Theorem 2.18. $\mathsf{P} + \Pi_1^1\text{-}\mathsf{TRANS}$ *and* $\mathsf{P} + \Pi_1^0\text{-}\mathsf{TRANS}$ *do not prove* STP *or* LMP.

Proof. Immediate from [33, Cor. 4.6]. □

Secondly, in light of the failure of $\Pi_1^0\text{-}\mathsf{TRANS} \to \mathsf{STP}$, it is a natural question how strong the combination $\Pi_1^0\text{-}\mathsf{TRANS} + \mathsf{STP}$ is. As it turns out, we readily obtain $\mathsf{ATR}^{\mathrm{st}}$ from $\Pi_1^0\text{-}\mathsf{TRANS} + \mathsf{STP}$. The same theorem for LMP fails.

Theorem 2.19. *The system* $\mathsf{P}_0 + \Pi_1^0\text{-}\mathsf{TRANS} + \mathsf{STP}$ *proves* $\mathsf{ATR}_0^{\mathrm{st}}$ *while* $\mathsf{P} + \Pi_1^0\text{-}\mathsf{TRANS} + \mathsf{LMP}$ *does not.*

Proof. Immediate from [33, Theorems 6.3 and 6.4]. □

Note that WKL and WWKL (and hence STP and LMP) are 'very close' in the sense that there is nothing between them in the RM zoo ([7]) or the Weihrauch degrees ([3]).

Theorem 2.20. *The system* $\mathsf{P} + \Pi_1^0\text{-}\mathsf{TRANS} + \mathsf{LMP}$ *does not prove* STP.

Proof. Immediate from Theorem 2.19. □

Finally, we often use this theorem without mention.

Theorem 2.21. *If* RCA_0 *proves* A, *then* P_0 *proves* A^{st}.

Proof. One readily verifies that P_0 proves the axioms of RCA_0 relative to 'st'. □

2.5. **Open questions.** The above listed theorems from [33] give rise to the following open questions. They will be answered in this paper.

First of all, in light of Theorem 2.11, it is a natural question how strong $\Theta + \mu^2$ is compared to well-known functionals. We show in Section 3.2 that S^2 is not computable from $\Theta + \mu^2$. In Section 3.3, we also provide a direct proof (not involving Nonstandard Analysis) of the fact that $\Theta + \mu^2$ computes a realiser for ATR_0.

Secondly, in light of Theorem 2.18, it is a natural question 'how high' Π_1^1-$\mathsf{TRANS}+$ STP actually goes. We show in Section 4.1 that the latter combination exists at the level of Π_2^1-CA_0, i.e. strictly stronger than Π_1^1-CA_0 and Π_1^1-TRANS. As a result, Π_1^1-$\mathsf{CA}_0^\omega + \mathsf{QF}\text{-}\mathsf{AC}^{2,1} + \mathsf{HBU}$ proves the Π_3^1-consequences of Π_2^1-CA_0.

Thirdly, in light of Theorem 2.14 and 2.19, it is a natural question whether weak fan functionals carry non-trivial strength. The answer is negative, in the following sense: we will identify a weak fan functional Λ_1 and show in Section 3.4 that $\Lambda_1 + \mu^2$ computes the same objects as μ^2. This shows that we cannot in general compute a special fan functional from a weak one. This provides mathematical evidence for the intuition that compactness up to measure is strictly weaker than full compactness.

Fourth, in light of Theorem 2.19, it is a natural question whether our results somehow generalise to Schweber's generalisation of ATR_0 in third-order arithmetic [51, 52]. We obtain such a generalisation for Theorem 2.19 in Section 4.2.

2.6. **Equivalent definitions.** We show that the definition of the special and weak fan functionals from Section 2.1 is equivalent to the original definition from [44].

The following definition for special fan functionals was used in [44]. We reserve the variable 'T^1' for trees and denote by '$T^1 \le 1$' that T is a binary tree.

Definition 2.22. The formula $\mathsf{SCF}(\nu)$ is as follows for $\nu^{(2 \to (0 \times 1^*))}$:

$$(\forall g^2, T^1 \le 1)\big[(\forall \alpha \in \nu(g)(2))(\overline{\alpha}g(\alpha) \notin T) \to (\forall \beta \le 1)(\exists i \le \nu(g)(1))(\overline{\beta}i \notin T)\big].$$

The provenance of the name of the specification '$\mathsf{SFF}(\Theta)$' for the <u>s</u>pecial <u>f</u>an <u>f</u>unctional is obvious. Similarly, $\mathsf{SCF}(\eta)$ was initially (and incorrectly) believed to be a special case of the (classical) fan functional, explaining its name. We now have the following theorem.

Theorem 2.23. *There are terms s, t of Gödel's T of lowest level such that*

$$(\forall \Theta)(\mathsf{SFF}(\Theta) \to \mathsf{SCF}(t(\Theta))) \wedge (\forall \nu)(\mathsf{SCF}(\nu) \to \mathsf{SFF}(s(\nu))). \tag{2.17}$$

Proof. We first provide a proof based on Computability Theory. Define $s(\nu) := \lambda g.\nu(g)(2)$ and define $t(\Theta) := \lambda g.(\max\{g(\alpha) \mid \alpha \in \Theta(g)\} + 1, \Theta(g))$. Assume $\mathsf{SCF}(\nu)$ and for given g consider $\nu(g) = (n, \{\alpha_1, \ldots, \alpha_k\})$. If $\mathsf{SFF}(s(\nu))$ fails for g, there is a $\beta \le 1$ that is not in any $[\overline{\alpha_j}g(\alpha_j)]$. Let T be the tree of all sequences $\overline{\beta}m$. Then the antecedent in $\mathsf{SCF}(\nu)$ holds for g and this T, but not the conclusion.

Now assume $\mathsf{SFF}(\Theta)$ and let g be given. We have that $\Theta(g) = \{\alpha_1, \ldots, \alpha_k\}$ where $C = [\overline{\alpha_1}g(\alpha_1)] \cup \cdots \cup [\bar{\alpha}_k g(\alpha_k)]$. We must prove $\mathsf{SCF}(t(\Theta))$. Again we argue by contradiction. Let T be a binary tree such that there is a β with $\overline{\beta}i \in T$, where $i = (\max\{g(\alpha) \mid \alpha \in \Theta(g)\} + 1$, i.e. the conclusion in $SCF(t(\Theta))$ fails for this T. Then $\beta \in [\bar{\alpha}_j g(\alpha_j)]$ for some $1 \le j \le k$ and $\bar{\alpha}_j g(\alpha_j)$ will be a sub-sequence of $\overline{\beta}i$. Thus the assumption in $\mathsf{SCF}(t(\Theta))$ does not hold for this T either.

We also provide a proof based on Nonstandard Analysis. Following Theorem 2.6, P_0 proves that the normal form (2.12) is equivalent to the normal form

$$(\forall^{\mathrm{st}}G^2)(\exists^{\mathrm{st}}w^{1^*})(\forall f^1 \leq 1)(\exists g \in w)(f \in [\overline{g}G(g)]). \tag{2.18}$$

Since standard functionals provide standard output for standard input, $(\exists^{\mathrm{st}}\Theta)\mathsf{SFF}(\Theta)$ implies (2.18). Hence, P_0 also proves the following:

$$(\forall^{\mathrm{st}}\Theta)[\mathsf{SFF}(\Theta) \rightarrow (2.12)]. \tag{2.19}$$

Now bring outside the standard quantifiers in the consequent of (2.19) and apply term extraction as in Corollary 2.4 to obtain the first conjunct of (2.17). The second conjunct follows in the same way. $\square$

We now discuss the definition of the weak fan functionals similar to Definition 2.22. We first introduce *weak weak König's lemma*.

Definition 2.24. [Weak weak König's lemma]
  (i) For $T \leq 1$, define $L_n(T) := \frac{|\{\sigma \in T : |\sigma| = n\}|}{2^n}$.
  (ii) For $T \leq 1$, define[5] '$\mu(T) >_{\mathbb{R}} 0$' as '$\lim_{n\to\infty} L_n(T) >_{\mathbb{R}} 0$'.
  (iii) We define WWKL as $(\forall T \leq 1)\big[\mu(T) >_{\mathbb{R}} 0 \rightarrow (\exists \beta \leq 1)(\forall m)(\overline{\beta}m \in T)\big]$.

As noted right after Definition 2.1, special fan functionals intuitively provide a finite sub-cover on input an uncountable cover of $2^{\mathbb{N}}$. Similarly, weak fan functionals provide an enumerated set of neighbourhoods covering a set of measure one. Again similar to the special ones, the weak fan functionals originate from a weak version of the nonstandard compactness of Cantor space, as discussed in Section 2.2.

Definition 2.25. The formula $\mathsf{WCF}(\eta)$ is as follows for $\eta^{(2\to(1\times 1^*))}$:

$$(\forall k^0, g^2, T^1 \leq 1)\big[(\forall \alpha \in \eta(g,k)(2))(\overline{\alpha}g(\alpha) \notin T) \rightarrow L_{\eta(g,k)(1)}(T) \leq \tfrac{1}{2^k}\big].$$

In contrast to ν, η only outputs (via the function $\lambda k.\eta(g,k)(1)$) a modulus for $\mu(T) = 0$ rather than a finite upper bound for T. The antecedent in the definition of η is similar to that of ν: a finite sequence of paths not in T is provided (via $\eta(g,k)(2)$). Thus, there is a trivial term of Gödel's T computing η in terms of ν.

Similar to Theorem 2.17, we have the following equivalence.

Theorem 2.26. *There are terms s, t of Gödel's T of lowest level such that*

$$(\forall \Lambda)(\mathsf{WFF}(\Lambda) \rightarrow \mathsf{WCF}(t(\Lambda))) \wedge (\forall \eta)(\mathsf{WCF}(\eta) \rightarrow \mathsf{SFF}(s(\eta))). \tag{2.20}$$

The first proof of Theorem 2.23 is easily adjusted to a proof of Theorem 2.26.

3. Uniform computability for Θ, Λ, and μ^2

In this section, we investigate uniform Kleene-computability for respectively special and weak fan functionals Θ and Λ, combined with Feferman's μ. In Section 3.1 we discuss some preliminary results and notation. In Section 3.2, we show that *only* hyperarithmetical functions can be *uniformly* computed by Θ and μ; as a result, the latter combination does not compute the Suslin functional. In Section 3.3, we provide a *direct* proof that ATR_0 can be obtained from Θ and μ^2, which was established *indirectly* (using term extraction from Nonstandard Analysis) in [33, §6].

---

[5]Note that a statement of the form '$\lim_{n\to\infty} a_n >_{\mathbb{R}} b$' always makes sense as a formula of second-order arithmetic, namely $(\exists N^0)(\exists k^0)(\forall n^0 \geq N)(a_n >_{\mathbb{R}} b + \frac{1}{2^k})$, even if limit at hand cannot be proved to exist in a weak system, like the base theory RCA_0.

Thus, the combination Θ plus μ^2 can compute *non-hyperarithmetical* functions, *but only non-uniformly*. By contrast, in Section 3.4, we construct Λ_1, a weak fan functional such that only hyperarithmetical functions are computable in Λ_1 and μ. As a consequence, special fan functionals are in general not computable from a weak fan functional Λ combined with μ.

3.1. Preliminaries.

3.1.1. *Introduction.* In this section, we introduce the Kleene schemes S1-S9 and consider some minor modifications due to the need for notational simplicity. We are primarily interested in the computational power of special fan functionals Θ or weak fan functionals Λ, in conjunction with Feferman's μ. We establish our results with respect to full Kleene computability. For this, it does not matter if we consider Kleene's $\exists^2$ or Feferman's μ, but in case we restrict ourselves to primitive recursion, μ is no longer computable in $\exists^2$. Thus, for studying the computational power of sub-classes of S1-S9 like fragments of Gödel's T, it is better to use μ.

For the reader unacquainted with (higher-order) computability theory, we point to some well-known facts that we will use without further reference:

(i) For subsets of $\mathbb{N}$ or $\mathbb{N}^{\mathbb{N}}$, the *hyperarithmetical* sets are exactly those computable in μ, or equivalently in $\exists^2$, and exactly the Δ_1^1-sets.

(ii) The Π_1^1-sets are exactly the sets *semi-computable* in μ (or $\exists^2$), i.e. the domains of functions partially computable in μ.

(iii) The ordinal ω_1^{CK} ('CK' for *Church-Kleene*) is the least ordinal without a computable code. Gödel's $L_{\omega_1^{\mathsf{CK}}}$, the fragment of the universe of the constructible sets up to ω_1^{CK}, is the least Σ_1-admissible structure[6].

3.1.2. *The functionals Θ and Λ.* We will investigate uniform Kleene-computability for respectively Θ and Λ combined with μ. We now provide suitable alternative definitions of these fan functionals to be used below.

According to the specification $\mathsf{SFF}(\Theta)$, Θ is a functional of type $2 \to 1^*$ where for each F, the set of neighbourhoods $C_{\bar{g}(F(g))}$, with $g \in \Theta(F)$, is a cover of the Cantor space. For adjustment to the Kleene schemes, it is better to use an alternative presentation, coding a finite sequence from C into one, as follows.

In this section, we let $\Theta(F)$ be an element of Cantor space that is not constant zero. Each such object f will code a finite sequence $\langle g_1, \ldots, g_k \rangle$ of binary functions by letting k be the least positive number such that $f(k-1) = 1$, and then decode $g(n) = f(n + k)$ into k elements using the standard k-partition of $\mathbb{N}$, i.e. $g_i(m) = g(m \cdot k + i - 1)$. When s is a finite binary sequence, we also use C_s to denote the corresponding basic neighbourhood in C, essentially meaning the same as the formal expression $[s]$. We will write $\Theta(F) = \langle g_1, \ldots, g_k \rangle$ and we will assume that Θ satisfies that for all F, $\{C_{\bar{g}_i(F(g_i))} \mid i = 1, \ldots, k\}$ is a cover of Cantor space. The latter is equivalent to stating that for some $n \in \mathbb{N}$ and for all binary sequences s of length n there is some i such that $\bar{g}_i(F(g_i))$ is an initial segment of s.

Similarly, according to the specification $\mathsf{WFF}(\Lambda)$, $\Lambda(F)(k)$ is a finite sequence $\langle f_1, \ldots, f_n \rangle$ from Cantor space such that $\mathbf{m}(\bigcup_{i=1}^n C_{\bar{f}_i(F(f_i))}) \geq 1 - \frac{1}{2^k}$, where $\mathbf{m}$

---

[6]A structure is Σ_1-admissible if it satisfies the Kripke-Platek axioms Δ_1-comprehension and Σ_1-replacement. We say that an ordinal α is admissible if the corresponding fragment of L is admissible.

denotes the standard product measure on Cantor space C. When studying aspects of computability relative to Λ and μ, we may equivalently let $\Lambda(F)$ be a sequence $(f) = (f_i)_{i \in \mathbb{N}}$ such that $\mathbf{m}(\bigcup_{i \in \mathbb{N}} C_{\bar{f}_i(F(f_i))}) = 1$. For notational reasons, this is the form for Λ we will use in this section.

3.1.3. *The Kleene Schemes.* Turing's famous model of computability ([59]) is restricted to inputs of types zero and oracles of type one. By way of generalisation, Kleene introduces computations taking sequences $\vec{\Phi}$ of higher order functionals Φ of pure types as arguments ([15]). In particular, via the schemes S1-S9, that are clauses in a grand monotone inductive definition, he defined the relation $\{e\}(\vec{\Phi}) = a$, i.e. the e-th (Kleene) computation with input $\vec{\Phi}$ terminates with output $a \in \mathbb{N}$.

For the purpose of this section, we will introduce the Kleene schemes S1-S9 with some minor modifications, motivated by he following:

(i) In all our computations, at most one functional of type 2 is used as an argument, namely Feferman's μ.

(ii) The scheme S8 for functional application was originally designed for functionals of pure type. However, special fan functionals are of mixed type $(\mathbb{N}^{\mathbb{N}} \to \mathbb{N}) \to (\mathbb{N} \to \mathbb{N})$ while weak fan functionals are of type $(\mathbb{N}^{\mathbb{N}} \to \mathbb{N}) \to (\mathbb{N} \to (\mathbb{N} \to \mathbb{N}))$.

Instead of coding Θ and Λ as objects of pure type 3, we modify the schemes S1-S9 so that they make sense for the one argument μ of type 2 and for any functionals Θ, and later Λ, of the relevant mixed types. The only motivation for this adjustment to mixed types is readability: we will let the special and weak fan functionals appear directly in the schemes, and not in coded form. It is a matter of unpleasant routine to show that this modification yields the same notion of computation as Kleene's original schemes via the standard reductions to pure types. In [20, Section 5.1.3], Kleene's notion of computation is extended to all finite types via some form of λ-calculus, but we prefer not to introduce the general machinery here.

Assume that the functional Θ is of the specified type. Let $\vec{g}$ be a sequence of functions and $\vec{b}$ be a sequence of numbers. We now define the relation $\{e\}(\Theta, \mu, \vec{g}, \vec{b}) = a$ by induction as follows.

Definition 3.1 (Modified Kleene S1-S9)**.**

(S1) $\{\langle 1 \rangle\}(\Theta, \mu, \vec{g}, a, \vec{b}) = a + 1$

(S2) $\{\langle 2, a \rangle\}(\Theta, \mu, \vec{g}, \vec{b}) = a$

(S3) $\{\langle 3 \rangle\}(\Theta, \mu, \vec{g}, a, \vec{b}) = a$

(S4) If $e = \langle 4, e_1, e_2 \rangle$ and for some b we have that

(i) $\{e_1\}(\Theta, \mu, \vec{g}, \vec{b}) = b$

(ii) $\{e_2\}(\Theta, \mu, \vec{g}, b, \vec{b}) = a$

then $\{e\}(\Theta, \mu, \vec{g}, \vec{b}) = a$

(S5) If $e = \langle 5, e_1, e_2 \rangle$ then (with the obvious interpretation, in analogy with S4)

(i) $\{e\}(\Theta, \mu, \vec{g}, 0, \vec{b}) = \{e_1\}(\Theta, \mu, \vec{g}, \vec{b})$

(ii) $\{e\}(\Theta, \mu, \vec{g}, a + 1, \vec{b}) = \{e_2\}(\Theta, \mu, \vec{g}, a, \{e\}(\Theta, \mu, \vec{g}, a, \vec{b}), \vec{b})$

(S6) Let $\vec{g} = (g_1, \ldots, g_k)$, $\vec{b} = (b_1, \ldots, b_m)$ and let τ_1, τ_2 be permutations of $\{1, \ldots, k\}$ and $\{1, \ldots, m\}$ respectively. If $e = \langle 6, e_1, \tau_1, \tau_2 \rangle$ then

$$\{e\}(\Theta, \mu, g_1, \ldots, g_k, a_1, \ldots, a_n) = \{e_1\}(\Theta, \mu, g_{\tau_1(1)}, \ldots, g_{\tau_1(k)}, b_{\tau_2(1)}, \ldots, b_{\tau_2(m)})$$

(S7) $\{\langle 7 \rangle\}(\Theta, \mu, g, \vec{g}, b, \vec{b}) = g(b)$

(S8.1) If $e = \langle 8, 1, e_1 \rangle$ and $\{e_1\}(\Theta, \mu, \vec{g}, a, \vec{b})$ is defined for all $a \in \mathbb{N}$ then

    (i) $\{e\}(\Theta, \mu, \vec{g}, \vec{b}) = 0$ if $\{e_1\}(\Theta, \mu, \vec{g}, a, \vec{b}) = 0$ for all a

    (ii) $\{e\}(\Theta, \mu, \vec{g}, \vec{b}) = a$ for the least a such that $\{e_1\}(\Theta, \mu, \vec{g}, a, \vec{b}) > 0$ otherwise

(S8.2) If $e = \langle 8, 2, e_1 \rangle$, let $F(g) = \{e_1\}(\Theta, \mu, g, \vec{g}, \vec{b})$. If F is total, we let $\{e\}(\Theta, \mu, \vec{g}, a, \vec{b}) = \Theta(F)(a)$.

(S9) If $e = \langle 9, i, j \rangle$, $i \le k$ and $j \le m$, then

$$\{e\}(\Theta, \mu, g_1, \ldots, g_k, d, b_1, \ldots, b_m) = \{d\}(\Theta, \mu, g_1, \ldots, g_i, b_1, \ldots, b_j)$$

If we leave out S9 in the previous definition, we have the schemes for Kleene primitive recursion. Furthermore, a definition of the relation $\{e\}(\Lambda, \mu, \vec{g}, \vec{b}) = a$, where Λ is a weak fan functional, is obtained by replacing Θ with Λ everywhere in S1 - S7, S9 and S8.1, and replacing S8.2 with the following formula:

(S8.3) If $e = \langle 8, 3, e_1 \rangle$, put $F(g) = \{e_1\}(\Lambda, \mu, g, \vec{g}, \vec{b})$. If F is total, define the value $\{e\}(\Lambda, \mu, \vec{g}, i, a, \vec{b})$ as $\Lambda(F)(i)(a)$.

All these schemes are viewed as clauses in a strictly positive inductive definition. If we leave out S9, then the definition may be viewed as a recursion on e. The set of indices, together with the relevant arities, can then be defined by standard primitive recursion over $\mathbb{N}$. Moreover, in this case all 'computations' will terminate, as partiality is only introduced via S9.

3.2. **Uniform computability in Θ.** In this section, we will introduce the notion of a Θ-*structure* (see Definition 3.5) and use the associated model theory to prove two crucial theorems (Theorems 3.2 and 3.3) regarding computability in μ and Θ. As a corollary, we obtain that $\Theta + \mu$ does not compute S^2. The proof in this section can be viewed as an elaboration on the proof of [20, Theorem 5.2.25].

First of all, as to notation, recall that ω_1^f is the least ordinal not represented by any well-ordering Turing-computable in f (see [43, X.2.9]). Also, throughout this section, the quantifier '$\forall \Theta$' is to be understood as 'for all special fan functionals Θ', i.e. $(\forall \Theta)(\mathsf{SFF}(\Theta) \to \ldots)$, which we omit for reasons of space.

Theorem 3.2. *There is a special fan functional Θ such that for all functions f computable in Θ and μ we have that $\omega_1^f = \omega_1^{\mathrm{CK}}$.*

Theorem 3.3. *The set $\{(e, \vec{y}, a) \mid \forall \Theta. \{e\}(\Theta, \mu, \vec{y}) = a\}$ is Π_1^1, where $\vec{y}$ ranges over all finite sequences of non-negative integers.*

The following corollary implies that Θ and μ cannot uniformly compute S^2.

Corollary 3.4. *Let f be a function such that for some e, $\{e\}(\Theta, \mu, n) = f(n)$ for all n and all special fan functionals Θ. Then f is hyperarithmetical.*

We could, in Theorem 3.3, let $\vec{y}$ range over all sequences of objects of type zero and one, but we have not found any use for this observation. The proof of Theorem 3.3 will essentially be an application of the Löwenheim-Skolem theorem, establishing the fact that the following statements are equivalent:

    (i) For all Θ, $\{e\}(\Theta, \mu, \vec{g}, \vec{b}) = a$.

    (ii) For all countable models $\mathcal{M}$ containing $\vec{g}$ and a special fan functional $\Theta_{\mathcal{M}}$ (in the sense of the model as indicated), $\mathcal{M} \models \{e\}(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a$.

We must, however, show some care in what we mean by 'a model' and what we then mean by '$\mathcal{M} \models \{e\}(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a$'. For instance, we cannot use the usual inductive definition involved in Kleene computability directly, because the least fixed point of the Kleene schemes, even when restricted to a countable structure, is Π_1^1 itself. Moreover, the Löwenheim-Skolem argument does not work for second-order concepts, so we need to replace Kleene's definition with something first-order. It turns out that it suffices to consider *all* fixed points of the Kleene schemes. Also, the proof of Theorem 3.3 yields Theorem 3.2 'almost for free'.

We introduce the notion of a Θ-structure as follows.

Definition 3.5. A Θ-*structure* is a tuple $\mathcal{M} = \langle \mathbb{N}, M_1, M_2, \Theta_{\mathcal{M}}, \mu, R \rangle$ such that

(i) M_1 is a set of functions $f : \mathbb{N} \to \mathbb{N}$ and M_2 is a set of functions $G : M_1 \to \mathbb{N}$.
(ii) $\mu \in M_2$ satisfies the usual definition of μ.
(iii) $\Theta_{\mathcal{M}} : M_2 \to \{0,1\}^{\mathbb{N}}$ and satisfies the modified $\mathsf{SFF}(\Theta)$ relative to M_1, M_2, see Section 3.1.2.
(iv) R stands for a relation $[e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a$, where $\vec{g}$ is a finite sequence from M_1 and $\vec{b}$ is a finite sequence from $\mathbb{N}$, that satisfies:
   (a) For each $\vec{g}, \vec{b}$ there is at most one a such that $[e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a$.
   (b) If for some $\vec{g}, \vec{b}, e$, we have $(\forall b \in \mathbb{N})(\exists a)([e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, b, \vec{b}) = a)$, then there is an $f \in M_1$ such that $(\forall b \in \mathbb{N})([e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, b, \vec{b}) = f(b))$.
   (c) If for some $\vec{g}, \vec{b}, e$ we have $(\forall g \in M_1)(\exists a)([e]_R(\Theta_{\mathcal{M}}, \mu, g, \vec{g}, \vec{b}) = a)$, then there is a $G \in M_2$ such that $(\forall g \in M_1)[e]_R(\Theta_{\mathcal{M}}, \mu, g, \vec{g}, \vec{b}) = G(g)$.
   (d) The relation R is a fixed point of the Kleene schemes from Definition 3.1 interpreted over $\mathcal{M}$.

We will not distinguish in notation between μ in the structure $\mathcal{M}$ and μ in the full universe. For the below proofs, we need to code countable Θ-structures as objects of type 1. Clearly, the set of codes for countable Θ-structures will be arithmetical:

Definition 3.6. Let $\mathcal{M} = \langle \mathbb{N}, M_1, M_2, \mu, \Theta_{\mathcal{M}}, R \rangle$ be a countable Θ-structure. A *code* for $\mathcal{M}$ is a function $f = \langle f_1, f_2, f_3, f_4 \rangle$ such that

(i) $f_1 = \langle f_{1,i} \rangle_{i \in \mathbb{N}}$ enumerates M_1 in a 1-1-way.
(ii) Let $f_2 = \langle f_{2,j} \rangle_{j \in \mathbb{N}}$ and let $F_j(f_{1,i}) = f_{2,j}(i)$. Then $\{F_j\}_{j \in \mathbb{N}}$ enumerates M_2 in a 1-1-way.
(iii) $f_3(\langle j, a \rangle) = \Theta_{\mathcal{M}}(F_j)(a)$ for all j and a.
(iv) $f_4(\langle e, \langle i_1, \ldots, i_k \rangle, \langle b_1, \ldots, b_m \rangle, a \rangle) = 0$ if and only if $[e]_R(\Theta_{\mathcal{M}}, \mu, f_{i_1}, \ldots, f_{i_k}, b_1, \ldots, b_m) = a$.

It is essential for the below argument that the set of codes for Θ-structures is arithmetical (or at least hyperarithmetical). The crucial part here is the first-order definition of special fan functionals. The same result can be obtained for some other (classes of) functionals, but e.g. not for the Superjump or the Suslin functional. For those interested in such a generalisation, note that replacing SFF with another class of functionals Γ requires that one can relativise Γ to type structures $\mathcal{M}$.

Definition 3.7. Let $\mathcal{M}$ be a Θ-structure. An *extension* of $\Theta_{\mathcal{M}}$ is a special fan functional Θ_1 such that whenever F of type 2 is an extension of $G \in M_2$ then $\Theta_1(F) = \Theta_{\mathcal{M}}(G)$.

Lemma 3.8. *For any Θ-structure $\mathcal{M}$, the functional $\Theta_{\mathcal{M}}$ has an extension Θ_1.*

Proof. Let Θ_0 be any special fan functional, for instance the one constructed in [33, §5]. We define

$$\Theta_1(F) := \begin{cases} \Theta_{\mathcal{M}}(G) & \text{if } G \in M_2 \text{ and } F \text{ extends } G \\ \Theta_0(F) & \text{if there is no such } G \in M_2 \end{cases}.$$

The definition of special fan functionals does not require any connection between the values of $\Theta(F_1)$ and $\Theta(F_2)$ when $F_1 \neq F_2$: we have only specified how F and $\Theta(F)$ are related for each F. This relation will hold point-wise for each $(F, \Theta_1(F))$ by construction, so Θ_1 will also be a special fan functional. $\square$

We could provide a similar construction and prove a similar lemma for other classes of type 3 functionals, but not for all. Actually, we would always be able to find *extensions* in a set-theoretical sense as above, but not necessarily in the class of functionals that we are interested in. The key property for us is that for a given F we specify, individually for that F, what an acceptable output of F will be in such a way that we only have to know F restricted to a countable (in this case, finite) set to justify that an alleged output is an acceptable one. If we, for instance, were interested in computations relative to $\exists^3$, we could not prove an extension lemma as above, since the constant zero in $\mathcal{M}$ will have extensions that are not constant zero, so the value of $\exists^3_{\mathcal{M}}$ cannot be preserved through extensions.

Even though the relation R does not have to represent the least fixed point of the Kleene schemes restricted to $\mathcal{M}$, we will see that it will contain this least fixed point as a sub-relation. In fact, we have the following lemma, where we only make use of extensions in general, not of the fact that we deal with special fan functionals.

Lemma 3.9. *Let $\mathcal{M} = \langle \mathbb{N}, M_1, M_2, \Theta_{\mathcal{M}}, \mu, R \rangle$ be a Θ-structure. Let Θ_1 be an extension of $\Theta_{\mathcal{M}}$ as above. Let $\vec{g}$ be a sequence from M_1 and $\vec{b}$ a sequence from $\mathbb{N}$. If $\{e\}(\Theta_1, \mu, \vec{g}, \vec{b}) = a$, then $[e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a$.*

Proof. We prove this by induction on the ordinal rank of the computation of $\{e\}(\Theta_1, \mu, \vec{g}, \vec{b})$. The proof will be given by cases following the schemes. For the schemes S1, S2, S3 and S7, the cases of initial computations, the claim follows directly from the assumption that R is a fixed point of the inductive operator whose least fixed point is the true set of terminating computations.

For the schemes S4 (composition), S5 (primitive recursion), S6 (permutation of arguments) and S9 (enumeration), the claim follows by the induction hypothesis and the assumption on R. This leaves us with the two special instances of S8:

- $\{e\}(\Theta_1, \mu, \vec{g}, \vec{b}) = \mu(\lambda x^0.\{d\}(\Theta_1, \mu, \vec{g}, x, \vec{b}))$. By the induction hypothesis and the closure properties of $\mathcal{M}$ we have

$$\lambda x.\{d\}(\Theta_1, \mu, \vec{g}, x, \vec{b}) = \lambda x.[d]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, x, \vec{b}) \in M_1,$$

  and the application of μ will yield the same result if we consider μ as an element of M_2 or as an element of full type 2. Then, since R is a fixed point of the Kleene computation operator, we have that

$$[e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = \{e\}(\Theta_1, \mu, \vec{g}, \vec{b}).$$

- $\{e\}(\Theta_1, \mu, m, \vec{g}, \vec{b}) = \Theta_1(\lambda g.\{d\}(\Theta_1, \mu, g, \vec{g}, \vec{b}))(m)$. By the induction hypothesis and the closure properties of $\mathcal{M}$, we have that

$$\lambda g \in M_1.\{d\}(\Theta_1, \mu, g, \vec{g}, \vec{b}) = \lambda g \in M_1.[d]_R(\Theta_{\mathcal{M}}, \mu, g, \vec{g}, \vec{b}) \in M_2. \qquad (3.1)$$

Let G^2 be the function defined by (3.1). Then $F = \lambda g \in \mathbb{N}^{\mathbb{N}}.\{e\}(\Theta_1, \mu, g, \vec{g}, \vec{b})$ is a total extension of G, so $\Theta_1(F) = \Theta_{\mathcal{M}}(G)$ by the assumption on Θ_1. The induction step then follows as above.

We have now treated all nine schemes, and the proof is done.   □

We need one more lemma as follows.

Lemma 3.10. *For each finite sequence $\vec{f}$ from $\mathbb{N}^{\mathbb{N}}$ and special fan functional Θ_1, there is a countable Θ-structure $\mathcal{M} = \langle \mathbb{N}, M_1, M_2, \mu, \Theta_{\mathcal{M}}, R \rangle$ with $\vec{f}$ in M_1 such that for all e, $\vec{g} \in M_1^*$, $\vec{b} \in \mathbb{N}^*$, and $a \in \mathbb{N}$, we have that*

$$[e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a \leftrightarrow \{e\}(\Theta_1, \mu, \vec{g}, \vec{b}) = a.$$

Proof. We define M_1 as a kind of Skolem hull, and we define M_2, $\Theta_{\mathcal{M}}$, and R explicitly from M_1 and Θ_1. We will need that Θ_1 is a special fan functional in order to show that $\mathcal{M}$ models that $\Theta_{\mathcal{M}}$ is a special fan functional, but the rest of the proof works for all type three objects.

Thus, let M_1 be countable such that

(i) Each f_i from $\vec{f}$ is in M_1
(ii) If g is computable in Θ_1, μ and a sequence $\vec{g}$ from M_1, then $g \in M_1$
(iii) If F is a partial functional of type 2 computable in Θ_1, μ, and some $\vec{g}$ from M_1, and there is some g for which $F(g)$ is undefined, then there is some $g \in M_1$ such that $F(g)$ is undefined. (This is the main Skolem hull part, and here we need the axiom of choice in a non-trivial way.)

We then let M_2 consist of all restrictions of F to M_1, where F is total and computable in Θ_1, μ and some $\vec{g}$ in M_1. If G is the restriction of F in this way, we put $\Theta_{\mathcal{M}}(G) := \Theta_1(F)$. We put $[e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a$ if and only if $\{e\}(\Theta_1, \mu, \vec{g}, \vec{b}) = a$ for $\vec{g}$ in M_1. Then (iii) will ensure that totality of functionals of type 2 is absolute for $\mathcal{M}$: If $\mathcal{M} \models \forall g \exists a [e]_R(\Theta, \mu, g, \vec{g}, \vec{b}) = a$, then F, defined by $F(g) = \{e\}(\Theta_1, \mu, \vec{g}, \vec{b})$, is total, and the restriction to $\mathcal{M}$ is in M_2.

By a similar argument, we observe that $\Theta_{\mathcal{M}}$ will be extensional: If $F_1 \neq F_2$, both are total and computable in Θ_1, μ and elements from M_1, then the partial functional F_3, where $F_3(g) = 0$ when $F_1(g) = F_2(g)$ and undefined otherwise, will also be computable in Θ_1, μ and elements from M_1, and by (iii), M_1 will contain a g such that $F_1(g) \neq F_2(g)$. Thus, the restriction operator will be 1-1, and $\Theta_{\mathcal{M}}$ is thus extensional, that is, well defined. Except for the construction of M_1, the construction of $\mathcal{M}$ is explicit. Moreover, if $F \in M_2$ and G is the unique extension of F computable in Θ_1, μ and elements from M_1, we have that $\Theta_1(G) \in M_1$, and that $\Theta_{\mathcal{M}}(F) = \Theta_1(G)$ codes a finite subset of M_1 that, together with G (or F) forms a finite cover of C, so $\Theta_{\mathcal{M}}$ will be a special fan functional from the point of view of $\mathcal{M}$. Thus $\mathcal{M}$ will satisfy the claim of the lemma.   □

Finally, we can prove Theorems 3.2 and 3.3 as follows.

Proof. (of Theorem 3.2) First of all, the functional Θ_0 is defined in [33, §5] and Lemma 3.10 implies that there is at least one countable Θ-structure $\mathcal{M}$, i.e. the set of codes for Θ-structures is hyperarithmetical and *non-empty*. By (essentially) the Gandy basis theorem ([43, III.1.4]), there is then a code f for a countable Θ-structure $\mathcal{M} = \langle \mathbb{N}, M_1, M_2, \mu, \Theta_{\mathcal{M}}, R \rangle$ such that $\omega_1^f = \omega_1^{\mathsf{CK}}$. By Lemma 3.8, $\Theta_{\mathcal{M}}$

has an extension Θ_1, and by Lemma 3.9, all functions g computable in Θ_1 and μ are elements of M_1, and thus Turing computable in f. Then also $\omega_1^g = \omega_1^{\mathsf{CK}}$.   □

The following provides a proof for Theorem 3.3.

Proof. By Lemmas 3.8, 3.9 and 3.10, the following are equivalent, given $e, \vec{g}, \vec{b}, a$:

(i) For all special fan functionals Θ, we have that $\{e\}(\Theta, \mu, \vec{g}, \vec{b}) = a$

(ii) For all countable Θ-structures $\mathcal{M} = \langle \mathbb{N}, M_1, M_2, \mu, \Theta_{\mathcal{M}}, R \rangle$, we have that $[e]_R(\Theta_{\mathcal{M}}, \mu, \vec{g}, \vec{b}) = a$.

Via coding, the relation in (ii) is Π_1^1, so the relation in (i) must also be Π_1^1.   □

3.3. Beyond the hyperarithmetical via Θ and μ^2.

In this section, we provide a *direct* proof that the combination Θ and μ^2 computes a realiser for ATR_0.

We proved in [33] that there is no instance Θ such that all functions computable in Θ and μ are hyperarithmetical. We gave two proofs: one by a direct construction of a hyperarithmetical functional F such that $\Theta(F)$ can never be contained in the hyperarithmetical functions, and one by applying term extraction to

$$\mathsf{P}_0 \vdash \Pi_1^0\text{-}\mathsf{TRANS} + \mathsf{STP} \to [\mathsf{ATR}_0]^{\mathrm{st}}$$

which (indirectly) yields a realiser for ATR_0 in terms of Θ and μ^2. There are thus two proofs of essentially the same result, one explicit construction where we do not analyse the logical strength needed and one indirect, via term extraction, where the underlying logic is explicit. We consider both approaches to be of value.

In a nutshell, the aim of this section is to prove (inside ACA_0) that ATR_0 follows from the *Arithmetical Compactness of C*, defined as follows.

Definition 3.11 (Arithmetical Compactness of C). *For any arithmetically defined $F : C \to \mathbb{N}$, where we allow function parameters, there are $f_1, \ldots, f_n \in C$ such that*

$$C \subseteq C_{\bar{f}_1(F(f_1))} \cup \cdots \cup C_{\bar{f}_n(F(f_n))}.$$

With the exception that we have used the symbol 'Θ' for other purposes (namely to denote a special fan functional), we mostly adopt Simpson's notation regarding ATR_0 from [54, V.2], namely as follows.

Notation 3.12. Let $\Gamma(n, X, Z)$ be an arithmetical formula, inducing the operator

$$\hat{\Gamma}(X, Z) = \{n \mid \Gamma(n, X, Z)\}$$

seen as an inductive operator in the first set variable X. We assume $A \subseteq \mathbb{N}$ and let $<_A$ be a total ordering of A. We use A and $<_A$ as hidden parameters, and when using the variable Y, we implicitly assume that $Y \subseteq \mathbb{N}^2$. We define $Y_a := \{n \mid (a, n) \in Y\}$ and $Y^a := \{(b, m) \mid (b, m) \in Y \wedge b <_A a\}$ for $a \in A$. Finally, $H(Y, Z)$ is the arithmetical statement $(\forall a \in A)(Y_a = \hat{\Gamma}(Y^a, Z))$.

Theorem 3.13. *Given Γ as above, there is an arithmetical function G^2 such that if $F(g) = G(g, A, <_A, Z)$ (g varies over C) and $g_1, \ldots, g_n$ are as in* Arithmetical Compactness *for F, then we can construct (uniformly arithmetically in Z, A, $<_A$ and $g_1, \ldots, g_n$) a pair (Y, h) such that either $H(Y, Z)$ or $h : \mathbb{N} \to \mathbb{N}$ is a strictly $<_A$-descending sequence in A. The verification can be formalised in ACA_0^ω.*

Proof. Given g, we put $Y[g] := \{(b, k) \mid g(\langle b, k \rangle) = 0\}$. We now define G as follows: the number $G(g, A, <_A, Z)$ is defined to be

(i) 0 if $H(Y[g], Z)$ or there is no $<_A$-minmal a such that $(Y[g])_a \neq \hat{\Gamma}((Y[g])^a, Z)$.

(ii) $\langle a, k \rangle + 1$ if a is $<_A$ minimal such that $(Y[g])_a \neq \hat{\Gamma}((Y[g])^a, Z)$ and k is the least integer in the symmetric difference of $(Y[g])_a$ and $\hat{\Gamma}((Y[g])^a, Z)$.

Let $F(g) = G(g, A, <_A, Z)$ and let $g_1, \ldots, g_n$ be such that $C = C_{\bar{g}_1(G(g_1))} \cup \cdots \cup C_{\bar{g}_n(G(g_n))}$. If for some i we have $F(g_i) = 0$, then either $H(Y[g_i], Z)$ or this is not the case since there is no $<_A$-minimal a such that $(Y[g_i])_a \neq \hat{\Gamma}((Y[g_i])^a, Z)$. We select the least such g_i in the lexicographical ordering on C. In the first case, we let $Y = Y[g_i]$ and h be the constant zero, and in the second case we may also let $Y = Y[g_i]$, but we combine μ-recursion and primitive recursion and let h be a strictly descending $<_A$ sequence of a's such that $(Y[g_i])_a \neq \hat{\Gamma}((Y[g_i])^a)$.

The other possibility is that $F(g_i) = \langle a_i, k_i \rangle + 1$ for $i = 1, \ldots, n$. If there are $i \neq j$ such that $Y[g_i] \cap A \times \mathbb{N} \neq Y[g_j] \cap A \times \mathbb{N}$ and there is no $<_A$-minimal a with $(Y[g_i])_a \neq (Y[g_j])_a$, we can extract an infinite descending sequence in A from this information. We will show that the absence of such i and j will lead to a contradiction. So assume that there is no such i and j. Without loss of generality, we may assume that $a_1 \leq_A a_2 \leq_A \cdots \leq_A a_n$. We make three observations:

(i) If $Y[g_i] \cap (A \times \mathbb{N}) = Y[g_j] \cap (A \times \mathbb{N})$, then $a_i = a_j$.

(ii) If $Y[g_i] \cap (A \times \mathbb{N}) \neq Y[g_j] \cap (A \times \mathbb{N})$ and a is the A-least number where $(Y[g_i])_a$ and $(Y[g_j])_a$ differ, then $\min_A\{a_i, a_j\} \leq_A a$.

(iii) Given i, if g is such that $(Y[g])^{a_i} = (Y[g_i])^{a_i}$ and $(Y[g])_{a_i} = \hat{\Gamma}((Y[g_i])^{a_i}, Z)$, then g is not covered by $C_{\bar{g}_i(F(g_i))}$. Moreover, if $a_i <_A a_j$, then g_j will satisfy this property of g.

It follows that if g is such that $(Y[g])^{a_n} = (Y[g_n])^{a_n}$ and $(Y[g])_{a_n} = \hat{\Gamma}((Y[g_n])^{a_n}, Z)$, then g is not in any of the sets $C_{\bar{g}_i(F(g_i))}$, so these sets do not form a cover. This is the desired contradiction.

It is easy to see that all steps here can be formalised.   □

This gives an alternative proof of the following corollary.

Corollary 3.14. *There is no special fan functional Θ that, together with μ, computes only hyperarithmetical functions.*

Proof. It is well established that there is no hyperarithmetical realiser for ATR_0, see e.g. the proof of V.2.6 in [54].   □

We also have the following corollary relativising the proof above.

Corollary 3.15. *There is an arithmetically defined function $F : C^2 \to \mathbb{N}$ such that for no special fan functional Θ, the function $F(x) = \Theta(\lambda y.F(x, y))$ is Borel.*

Proof. For $X \subset \mathbb{N}$, there is a total ordering computable in X that is not a well-ordering, but such that there is no descending sequence in the ordering hyperarithmetical in X. Hence, there is no realiser for ATR_0 hyperarithmetical in any X, i.e. no realiser that is Borel. Since we can obtain a realiser for ATR_0 by section-wise application of Θ to an arithmetical functional of two variables, we are done.   □

Finally, Hunter introduces a functional in [12, p. 23] that constitutes a 'uniform' version of ATR_0. This functional is computable from Θ plus μ, as follows

Corollary 3.16. *Uniformly primitive recursive in μ^2 and Θ^3 there is a functional $T : \mathbb{N}^{\mathbb{N}} \times (2^{\mathbb{N}} \to 2^{\mathbb{N}}) \to 2^{\mathbb{N}}$ such that when f^1 codes a well-ordering $<_f$, then $T(f, F)$ satisfies the following recursion equation for a in the domain of $<_f$:*

$$\{b : \langle b, a \rangle \in T(f, F)\} = F(\{\langle c, d \rangle \in T(f, F) : d <_f a\}).$$

Proof. In the proof of Theorem 3.13, note the fact that Γ is arithmetical is (only) used to prove that the defined functional G is arithmetical. The full proof therefore relativises to any F of the relevant type. $\qquad\square$

3.4. **Not beyond the hyperarithmetical via Λ and μ^2.** In this section, we introduce a functional Λ_1 of type $(C \to \mathbb{N}) \to (\mathbb{N} \to C)$, where $C = 2^{\mathbb{N}}$ is the Cantor space, with the following two properties:

(i) If $\Lambda_1(F) = \{f_i\}_{i \in \mathbb{N}}$, then $\bigcup_{i \in \mathbb{N}} C_{\bar{f}_i(F(f_i))}$ has measure 1.
(ii) Only hyperarithmetical functions are (S1-S9) computable in Λ_1 and μ^2.

Our motivation for introducing Λ_1 is to show that there is a weak fan functional in which no special fan functional is computable relative to μ^2. In [36], this result is linked to the RM of measure theory, the original Vitali Covering theorem ([60]) in particular, and it is also generalised to recursion relative to the Suslin functional.

As discussed in Section 3.1.2, item (i) just means that Λ_1 is a weak fan functional up to computational equivalence. The existence of a functional Λ satisfying item (i) follows from the existence of Θ, and we let Λ_0 be some fixed instance of Θ. We define Λ_1 in equation (3.4) below, namely in terms of Λ_0 and by specifying a different value for certain F. Since item (i) does not require any connection between $\Lambda_1(F)$ and $\Lambda_1(G)$ when $F \neq_2 G$, we have much freedom in constructing Λ_1. Of course, item (ii) puts some clear restrictions on how we can define Λ_1. For instance, if F is hyperarithmetical, i.e. computable in μ, we must have that $\Lambda_1(F)$ is hyperarithmetical. This can be arranged using basic measure theory and the Sacks-Tanaka theorems for measure-theoretic uniformity (see below). The next challenge is presented by 'iterated' outputs like for instance

$$\Lambda_1(\lambda f.\Lambda_1(\lambda g.F(f, g))(17)); \qquad\qquad (3.2)$$

these also need to be hyperarithmetical whenever F is. Again, basic measure theory and the Sacks-Tanaka machinery come to our rescue: as it turns out, except for a set of f's of measure zero, we can use the same value for $\Lambda_1(\lambda g \in CF(f, g))$ *independent* of f. However, we cannot expect to be able to use the same value as the output value in (3.2): the more involved Λ_1 is in a computation $\{e\}(\Lambda_1, \mu, \vec{b})$, the harder it is to find a hyperarithmetical output. Our guiding idea is that we may us the same hyperarithmetical value of $\Lambda(F)$ for almost all F computable at a certain countable level. We make this precise, as follows.

In the construction of Λ_1, we use the available machinery from measure theory and hyperarithmetical theory (i.e. the computability theory of μ), to construct a well-ordered sequence of *possible* values for Λ_1 indexed over the first non-computable ordinal ω_1^{CK} and (indirectly) a set X of measure 1 so that whenever F is computable in Λ_1 and elements from X, then we may let $\Lambda_1(F)$ be in that sequence. This may look circular, but in reality, Λ_1 and our sequence will be defined by a simultaneous transfinite recursion over ω_1^{CK}. This transfinite recursion is unfortunately (and unavoidably, we believe) a rather complex one.

Now, let us consider the machinery we need. First of all, we assume without mentioning that all sets and functions are measurable. Actually, we will only work with subsets of finite or countable products of the Cantor space C that are Σ_1^1 or Π_1^1 relative to objects of type 1, so measurability will not be an issue. The Cantor space C will have measure 1, so all products will have measure 1. We use $\mathbf{m}$ for the measure on all such product spaces. We will let "almost everywhere" mean that the property holds except possibly on a set of measure 0, which in our cases means that the property holds on a set of measure 1. We write 'a.a.' as short for 'almost all'. We will rely on two facts from measure theory, where all spaces are products E or D of the Cantor space C.

Proposition 3.17.
  (i) *A countable intersection of sets with measure 1 has measure 1.*
  (ii) *If $X \subset E \times D$ then $\mathbf{m}(X) = 1$ if and only if $\mathbf{m}(\{e \mid (e,d) \in X\}) = 1$ for a.a. $d \in D$ (if and only if $\mathbf{m}(\{d \mid (e,d) \in X\}) = 1$ for a.a. $e \in E$).*

These facts can be found in any standard textbook on measure theory. Item (ii) is actually a special case of Fubini's theorem for characteristic functions.

By convention, we denote infinite sequences of binary functions as $(f) := \{f_j\}_{j \in \mathbb{N}}$.

Definition 3.18. Let F be a partial function from C to $\mathbb{N}$ and let (f) be a sequence.
  (i) We say that (f) *suffices* for F if $F(f_j)$ is defined for all $j \in \mathbb{N}$ and $\mathbf{m}(\bigcup_{j \in \mathbb{N}} C_{\bar{f}_j(F(f_j))}) = 1$.
  (ii) We say that (f) *fails* F if $F(f_i)$ is undefined for some $i \in \mathbb{N}$.

Now, (f) suffices for F exactly when (f) can be an acceptable value of $\Lambda_1(G)$ for all total G extending F. In the next lemmas, we will make the following intuition precise: we can choose the same value $\Lambda_1(F)$ for large parameterised classes of Fs, and we have a lot of freedom in choosing this common value.

All the below arguments are elementary from the point of view of measure theory.

Lemma 3.19. *Let $F : C \to \mathbb{N}$ be a partial (measurable) functional with measurable domain.*
  (i) *If the domain of F has measure 1, then $\{(f) \mid (f)$ suffices for $F\}$ has measure 1.*
  (ii) *If the domain of F has measure less than 1, then $\{(f) \mid (f)$ fails $F\}$ has measure 1.*

Proof. Proof of item (i): for $k \in \mathbb{N}$, we will prove that the set of (f) such that $\mathbf{m}(\bigcup_{j \in \mathbb{N}} C_{\bar{f}_j(F(f_j))}) > 1 - 2^{-k}$, has measure 1. To this end, let n_k be so large that $\mathbf{m}(\{f \in C \mid F(f) < n_k\}) > 1 - 2^{-k}$. Let $s_{k,1}, \ldots, s_{k,m_k}$ be the binary sequences $s_{k,l}$ of length n_k such that $\mathbf{m}(\{f \in C_{s_{k,l}} \mid F(f) < n_k\}) > 0$. Let $r_{k,l}$ be this positive measure. Then $\mathbf{m}(\bigcup_{l=1}^{m_k} C_{s_{k,l}}) > 1 - 2^{-k}$ and for each $s_{k,l}$ the set of (f) such that for some f_j, $F(f_j) < n_k$ and f_j extends $s_{k,l}$, has measure 1. Indeed the probability of not satisfying this is $\prod_{j=0}^{\infty}(1 - r_{k,l}) = 0$. Since a finite intersection of sets of measure 1 still has measure 1, our claim follows; the previous generalises to countable intersections and item (i) holds.

Proof of item (ii): in this case, the probability that f_i is in the domain of F is smaller than 1 by a fixed value. Then the infinite product of the domain of F has measure 0. Thus, (f) fails F when (f) is in the complement of this product. $\square$

Lemma 3.20. *Let $F : C^2 \to \mathbb{N}$ be a partial, measurable functional defined on a measurable set and put $F_g(f) = F(f, g)$. Then the set of $(f) \in C^{\mathbb{N}}$ such that for a.a. $g \in C$ we have that*

    (i) *If the domain of F_g has measure 1, then (f) suffices for F_g*

    (ii) *If the domain of F_g has measure < 1 then (f) fails F_g*

has measure 1.

Proof. Let X be the set of $\langle g, (f) \rangle$ such that either F_g is defined on a set of measure 1 and (f) suffices for F_g or F_g is defined on a set of measure < 1 and (f) fails F_g. By Lemma 3.19, this set has measure 1, since for all g we have for almost all (f) that $\langle g, (f) \rangle \in X$. Then, by Proposition 3.17, the set of (f) such that $\langle g, (f) \rangle \in X$ for almost all g has measure 1, and we are done. $\qquad\square$

Since all (finite or countable) products of C we consider are isomorphic (with the exception of C^0), we will apply the previous lemma in other cases than for C^2 as well. For technical reasons, we shall need a strengthening of Lemma 3.20 as follows.

Definition 3.21. Let $\vec{c}$ be a non-repeating sequence from $\mathbb{N}$. We define $(f)_{\vec{c}}$ as the sequence of f_i indexed via $\vec{c}$.

Lemma 3.22. *Let $\vec{c}$ be a non-repeating sequence of length k' and let $X \subseteq C^{k'+k}$ have measure 1. Let $F : C \times C^{k'+k} \to \mathbb{N}$ be a partial functional that is measurable with a measurable domain and put $F_{\vec{h},\vec{g}}(g) := F(g, \vec{h}, \vec{g})$ for $\vec{h} \in C^{k'}$ and $\vec{g} \in C^k$. Then the set of pairs $\langle (f), \vec{g} \rangle$ such that $\langle (f)_{\vec{c}}, \vec{g} \rangle \in X$ and*

    (i) *if the domain of $F_{(f)_{\vec{c}},\vec{g}}$ has measure 1, then (f) suffices for $F_{(f)_{\vec{c}},\vec{g}}$, and*

    (ii) *if the domain of $F_{(f)_{\vec{c}},\vec{g}}$ has measure < 1, then (f) fails $F_{(f)_{\vec{c}},\vec{g}}$,*

has measure 1.

Proof. Combining Proposition 3.17.(ii) with the arguments of Lemmas 3.19 and 3.20, the lemma follows easily. The assumption that $\vec{c}$ is non-repeating is essential here, since otherwise the set of possible $(f)_{\vec{c}}$ will have measure 0 and not 1. $\qquad\square$

Now we have established the measure-theoretical lingo we need for the construction of Λ_1. In order to prove the main technical lemma, we also need some theorems from higher computability theory. We have formulated them in the form we need. For proofs, see [43, Sections IV.1-2 and Section X.4]. We will actually need some of these results in relativised forms, as follows.

Proposition 3.23.

    (i) *If $A \subset C$ is computable in f and μ via index e, then the relation $\mathbf{m}(A) = 1$ is decidable in μ, uniformly in f and e.*

    (ii) [Sacks,Tanaka] *If $A \subset C$ is hyperarithmetical and $\mathbf{m}(A) > 0$, then A contains a hyperarithmetical element.*

    (iii) [Gandy Selection] *If a Π_1^1-set of functions contains a hyperarithmetical element, we may find one, effectively in μ.*

    (iv) [Sacks, Tanaka] *The set of $g \in C$ such that $\omega_1^g = \omega_1^{\mathsf{CK}}$ has measure 1.*

Proof. The items from the theorem are proved as follows in [43]. Item (i) is Theorem IV.1.3. Item (ii) is Theorem IV.2.2. Item (iii) is proved as Theorem X.4.1 in a more general form. Item (iv) is Corollary IV.1.6. $\qquad\square$

We have established the general machinery needed below, and now start working towards the main result of his section.

Convention 3.24. From now on, we let '$\prec$' be a total, computable ordering of $\mathbb{N}$ such that the well-ordered initial segment has length ω_1^{CK}; $\prec$ may not be a well-ordering, but we will not actively use this fact. We let $W = W(\prec)$ be the elements in the well-ordered part, and for $i \in W$ we let α_i be the ordinal rank of i in $\prec$.

It is well-known that orderings as in Convention 3.24 exist. The set of computable total orderings that in addition are well-orderings, is complete Π_1^1; the set of computable total orderings without hyperarithmetical infinite descending sequences is Σ_1^1. Thus there is one ordering that is of the latter kind but that is not of the former kind. Such orderings are known as *computable pseudo-wellorderings*. The well-ordered initial segment is Π_1^1, but not Δ_1^1 or hyperarithmetical.

Let $[f]$ be a (double) sequence $\{(f_i)\}_{i \in W} = \{f_{i,j}\}_{i \in W, j \in \mathbb{N}}$ in $(W \times \mathbb{N}) \to (\mathbb{N} \to \{0, 1\})$. Each $[f]$ like this will define a *partial* approximation $\Lambda_{[f]}$ to a weak fan functional in the following sense:

Definition 3.25. Let $[f]$ be as above. For $F : C \to \mathbb{N}$, we define
   (i) $\Lambda_{[f]}(F) = (f_i)$ if $i \in W$, (f_i) is sufficient for F and no $(f_{i'})$ is sufficient for F for $i' \prec i$.
   (ii) $\Lambda_{[f]}(F)$ is undefined if there is no such $i \in W$.

Our aim is to construct $[f]$ is such a way that all functions computable in any total extension Λ of $\Lambda_{[f]}$ and μ^2 are hyperarithmetical. However, such a construction requires controlling the complexity of computations relative to any such extension. One obstacle is the requirement in Kleene's S8 that the input functional F must be total. We get around this obstacle by considering a more liberal interpretation of S8, so that it works for partial inputs as well, as long as they contain the relevant information. Such interpretations are well-established; see e.g. [20, §6.4].

Definition 3.26. We define the relation '$\{e\}_{[f]}(\Lambda_{[f]}, \mu, \vec{g}, \vec{b}) = a$' by transfinite recursion. We keep the schemes S1-S7, S8.1 and S9 from Definition 3.1, only adding $[f]$ as an index everywhere. We omit S8.2 and give a new interpretation of S8.3:

   (S8.3) If $e = \langle 8, 3, e_1 \rangle$, let $F(g) = \{e_1\}_{[f]}(\Lambda_{[f]}, \mu, g, \vec{g}, \vec{b})$. Then F is in general a partial function of type 2. Let $i \in W$ be the $\prec$-least number such that:
      (i) the value $F(f_{i',j})$ is defined for all $i' \preceq i$ and all $j \in \mathbb{N}$,
      (ii) the sequence (f_i) suffices for F. If there is one such i, then define
      $$\{e\}_{[f]}(\Lambda_{[f]}, \mu, \vec{g}, a, b, \vec{b}) := f_{i,a}(b); \text{ undefined otherwise.}$$

Since '$\{e\}_{[f]}(\Lambda_{[f]}, \mu, \vec{g}, \vec{b}) = a$' as in the previous definition is defined as the least fixed point of a positive inductive operator, each sequence $\langle e, \Lambda_{[f]}, \mu, \vec{g}, \vec{b}, a \rangle$ in the relation has an ordinal rank. Since we only require -even in the case of S8.3- a countable set of immediate sub-computations to terminate, the rank of any terminating computation modulo a given $[f]$ is countable, and actually an ordinal computable in $[f]$ and the argument list $\vec{g}$.

We will only apply this definition in the case where the map $i \mapsto (f_i)$ is a function that is partially Kleene-computable in μ^2. In this case, the partial function $\{e\}_{[f]}(\Lambda_{[f]}, \mu, \vec{g}, \vec{b})$ will be computable in μ^2 as well. Our goal is to construct $[f]$ in

such a way that for all indices e, input arguments $\vec{b}$ from $\mathbb{N}$, and total extensions Λ of $\Lambda_{|f|}$, we have that

$$\{e\}(\Lambda_{[f]}, \mu, \vec{b}) \simeq \{e\}_{[f]}(\Lambda_{[f]}, \mu, \vec{b}), \tag{3.3}$$

where '$\simeq$' means that both sides are undefined or both sides are defined and equal. If we succeed, we obviously have that all functions Kleene-computable in Λ will be hyperarithmetical when Λ is a total extension of $\Lambda_{|f|}$. We will not need that the total extension Λ itself is a weak fan functional for this argument.

After the construction of $[f]$, we will prove (3.3) by induction on the ordinal rank of the true Kleene-computation $\{e\}(\Lambda, \mu, \vec{b})$. In order to make this proof work, we have to take into account that there are sub-computations with arguments from C. We will see that it will be possible to construct $[f]$ such that we only have to consider argument sequences $\vec{g}$ of length k from a Σ_1^1-set X_k of measure 1. Let us now outline the construction, and what we attempt to achieve at each step:

(i) We will construct $[f]$ by defining (f_i) by recursion on $i \in W$

(ii) In parallel to defining (f_i), we will for each integer k construct a hyper-arithmetical set $X_{i,k} \subseteq C^k$ of measure 1 such that for any set of parameters $\vec{g} \in X_{i,k}$, integer parameters $\vec{b}$, and index e, we will have that the intended proof by induction will work for computations $\{e\}(\Lambda, \mu, \vec{g}, \vec{b})$ of ordinal rank bounded by α_i, the rank of i in $(W, \prec)$.

(iii) (f_i) will be chosen as a hyperarithmetical sequence that is sufficient for almost all functionals that are total on a set of measure 1 via computations strictly bounded by α_i, and fails almost all the others.

(iv) To verify the key Lemma 3.29, we have to consider inputs $\vec{g}$ together with inputs of the form $f_{i,j}$. Thus we will consider input sequences $\vec{h}, \vec{g}$ where the sequences $\vec{h}$ are 'specified' as certain $f_{i,j}$ and the sequences $\vec{g}$ will vary over sets of measure 1.

We will point out where the sequences $\vec{h}$ are needed in our technical argument. The underlying idea is that we may pick $\Lambda(F)$ 'at random' and the probability of success is 1. However, this random value must be random with respect to values of Λ obtained while computing F from Λ. Making this idea precise, the need arises to take previous values of Λ into account.

As a convention, when we write $\{e\}_{\mathsf{ind}}(\Lambda_{\mathsf{ind}}, \mu, \vec{g}, \vec{b})$, where 'ind' is any index, we assume without mentioning that the length of $\vec{g}$ fits the expression.

Convention 3.27. If (f_k) is a sequence $\{f_{k,j}\}_{j\in\mathbb{N}}$ for all $k \preceq i \in W$, then we write $[f]_i$ for $\{f_{k,j}\}_{k\preceq i, j\in\mathbb{N}}$. Similarly, if (f_k) is a sequence $\{f_{k,j}\}_{j\in\mathbb{N}}$ for all $k \prec i$, then we write $[f]_{\prec i}$ for $\{f_{k,j}\}_{k\prec i, j\in\mathbb{N}}$.

Our definition of $\Lambda_{[f]}$ readily generalises to $\Lambda_{[f]_i}$ and $\Lambda_{[f]_{\prec i}}$, and so does the recursive definition of $\{e\}_{\Lambda_{[f]}}(\Lambda_{[f]}, \mu, \vec{g}, \vec{b})$.

Convention 3.28. In the formulation of the next lemma, we make us of three kinds of inputs: integers, elements of the form $f_{i,j}$ that can be seen as parameters, and sequences $\vec{g}$ from C^k that can be seen as variables. As a convention, we order them $(\vec{h}, \vec{g}, \vec{b})$. There is no harm in this since we may always use S6 to permute inputs. In the proof, we shall introduce a fourth category $(f_i)_{\vec{c}}$ in the recursion

step, objects that may be in the $\vec{h}$-part at later stages, but whose values have not been decided before we select the one (f_i) we want to use.

There is a small twist to this notation: for our construction and argument it is important that the sequence $\vec{c}$ is non-repeating, but for our application we may want to consider computations where the same function is used in several locations in the list of arguments. Instead of building up an unbearable notation, we assume that we have one case for each way of distributing the arguments $(\vec{h}, (f_i)_{\vec{c}}, \vec{g}, \vec{b})$ as a list of inputs. Thus, each case we treat in the proof in theory covers countably many cases. We will inform the reader when we actually make use of this.

We have formulated our next item as a lemma, but it is in reality a combination of a construction by recursion and a verification of the key properties of this construction. We will refer to the details of the construction in later proofs.

Lemma 3.29. *By transfinite recursion on $i \in W$, we can construct $[f] = \{(f_i)\}_{i \in W}$ and sets $X_{i,k} \subseteq C^k$ of measure 1 (for each $k \in \mathbb{N}$ and $i \in W$) such that an alleged computation $\{e\}_{[f]_i}(\Lambda_{[f]_i}, \mu, \vec{h}, \vec{g}, \vec{b})$ will terminate whenever the parameters satisfy the following:*

(i) *$i \in W$ has norm α_i, e is a Kleene-index, $\vec{b} \in$ seq and $\vec{g} \in X_{i,k}$,*

(ii) *$\vec{h}$ is a sequence from $\{f_{i',j} \mid i' \preceq i \wedge j \in \mathbb{N}\}$*

(iii) *there is some extension $[f']$ of $[f]_i$ such that $\{e\}_{[f']}(\Lambda_{[f']}, \mu, \vec{h}, \vec{g}, \vec{b})\downarrow$ with a computation of ordinal rank at most α_i.*

Proof. We will show how to construct (f_i) and $X_{i,k}$ from $[f]_{\prec i}$ and $\{X_{i',l} \mid i' \prec i, l \in \mathbb{N}\}$. The key steps in our construction are:

(a) For each k, find a hyperanalytical set $Z_{i,k} \subseteq C^{\mathbb{N}} \times C^k$ with measure 1, such that the induction step works for all $\langle (f), \vec{g} \rangle \in Z_{i,k}$ if we use (f) as our (f_i).

(b) We let Y_i be the set of (f) such that for each $k \in \mathbb{N}$, we have $\mathbf{m}(\{\vec{g} \mid \langle (f), \vec{g} \rangle \in Z_{i,k}\}) = 1$. Then Y_i has measure 1 and is hyperarithmetical by Proposition 3.23.(i).

(c) We then select $(f_i) \in Y_i$ computably in μ by the Sacks-Tanaka basis theorem (see Proposition 3.23.(ii)) and Gandy Selection (see Proposition 3.23.(iii)) computably in μ.

(d) Finally, we define $X_{i,k} := \{\vec{g} \mid \langle (f_i), \vec{g} \rangle \in Z_{i,k}\}$.

The hard work will be to carry out step (a): the remaining steps then all follow by our general machinery.

Now assume that $[f]_{\prec i}$ and each $X_{i',l}$, for $i' \prec i$ and $l \in \mathbb{N}$, are constructed satisfying the claim of the lemma. We define $X_{\prec i,k} = \bigcap_{i' \prec i} X_{i',k}$, noting that if i_0 is the $\prec$-least integer, then $X_{\prec i_0,k}$ is C^k. The *induction hypothesis* is that $\mathbf{m}(X_{\prec i,k}) = 1$ and that for each e, each $\vec{b}$, each $\vec{h}$ from $\{f_{i',j} \mid i' \prec i \wedge j \in \mathbb{N}\}$, each $k \in \mathbb{N}$ and each $\vec{g} \in X_{\prec i,k}$ we have that if there is any extension $[f']$ of $[f]_{\prec i}$ such that $\{e\}_{[f']}(\Lambda_{[f']}, \mu, \vec{h}, \vec{g}, \vec{b})\downarrow$ with a computation of ordinal rank less than α_i, then $\{e\}_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \mu, \vec{h}, \vec{g}, \vec{b}) \downarrow$.

Since at the end, we use the recursion theorem for μ, we also assume that $X_{\prec i,k}$ is hyperartithmetical, with an index computable from μ, i and k.

Firstly, we construct sets of measure 1 dealing with each of the following cases:

$$\{e\}(\Lambda, \mu, \vec{h}, (f_i)_{\vec{c}}, \vec{g}, \vec{b}),$$

where e is a fixed Kleene-index, $\vec{b}$ is a fixed input of integers, $\vec{h}$ is as above , $\vec{c}$ is a sequence of length k' and $\vec{g} \in C^k$. Recall that each such case covers countably many cases by Convention 3.28. Then we let $Z_{i,k}$ be the intersection of the sets constructed for each of the cases. The purpose of $\vec{c}$ is to specify which elements in the sequence (f_i) we are about to construct, will be used as arguments in the computation *without specifying* (f_i) *itself*. All together, there are only *countably* many cases, so our set $Z_{i,k}$ will also have measure 1. The constructions are quite explicit and the induction hypothesis readily implies that $Z_{i,k}$ is hyperarithmetical.

We now show what to do in the two cases of composition and application of Λ; the rest of the cases are trivial, or they follow the pattern of 'S4 - composition'. We first treat the scheme S4 as follows: let $\vec{h}$ and $\vec{c}$ be as above and consider the case

$$\{e\}(\Lambda, \mu, \vec{h}, (f_i)_{\vec{c}}, \vec{g}, \vec{b}) = \{e_1\}(\Lambda, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, \{e_2\}(\Lambda, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, \vec{b}), \vec{b}).$$

We need to find a set of pairs $\langle (f), \vec{g} \rangle$ of measure 1 that guarantees that the induction step for this case goes through. Now, by Proposition 3.17.(ii) and the induction hypothesis, the set of $\langle (f), \vec{g} \rangle$ such that $(f)_{\vec{c}}, \vec{g} \in X_{\prec i, k'+k}$ has measure 1. Choose (f) and $\vec{g}$ in this set, and let $[f']$ be any extension of $[f]_{\prec i}(f)$, where we add the sequence (f) to the end of the double sequence $[f]_{\prec i}$. If we have

$$\{e\}_{[f']}(\Lambda_{[f']}, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, \vec{b}) =$$

$$\{e_1\}_{[f']}(\Lambda_{[f']}, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, \{e_2\}_{[f']}(\Lambda_{[f']}, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, \vec{b}), \vec{b}) = a$$

via a computation of ordinal rank at most α_i, then $\{e_2\}_{[f']}(\Lambda_{[f']}, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, \vec{b}) = c$ for some c, and also $\{e_1\}_{[f']}(\Lambda_{[f']}, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, c, \vec{b}) = a$, both with computational ranks *strictly* below α_i. Then, since $(f)_{\vec{c}}, \vec{g} \in X_{\prec i, k'+k}$, we can apply the induction hypothesis and conclude that

$$\{e_2\}_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, \vec{b}) = c \text{ and } \{e_1\}_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \mu, \vec{h}, (f)_{\vec{c}}, \vec{g}, c, \vec{b}) = a.$$

Thus, with any choice of (f) as (f_i) and $\vec{g}$ as above, we have

$$\{e\}_{[f]_i}(\Lambda_{[f]_i}, \mu, \vec{h}, (f_i)_{\vec{c}}, \vec{g}, \vec{b}) = a,$$

as required for this case.

We now turn to the cases with application of Λ, i.e. computations of the form

$$\Lambda(\lambda g.\{e\}(\Lambda, \mu, \vec{h}, g, (f)_{\vec{c}}, \vec{g}, \vec{b})).$$

As before, we see that the set of $\langle (f), g, \vec{g} \rangle$ such that $g, (f)_{\vec{c}}, \vec{g} \in X_{\prec i, 1+k'+k}$ has measure 1.

We now define $F_{(f), \vec{g}}(g) := \{e\}_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \mu, \vec{h}, g, (f)_{\vec{c}}, \vec{g}, \vec{b})$ provided this computation terminates with ordinal rank $< \alpha_i$. We claim that the following three sets all have measure 1:

   (i) The set of $\langle (f), \vec{g} \rangle$ such that

$$\mathbf{m}(\{g \mid (f)_{\vec{c}}, g, \vec{g} \in X_{\prec i, 1+k'+k}\}) = 1 \text{ and } (f)_{\vec{c}}, f_j, \vec{g} \in X_{\prec i, 1+k'+k} \text{ for all } j \in \mathbb{N}.$$

   (ii) The set of $\langle (f), \vec{g} \rangle$ such that $(f)_{\vec{c}}, \vec{g} \in X_{\prec i, k'+k}$.
   (iii) The set of $\langle (f), \vec{g} \rangle$ such that either (a) or (b) holds, as follows:
      (a) The domain of $F_{(f), \vec{g}}$ has measure 1 and (f) is sufficient for $F_{(f), \vec{g}}$.
      (b) The domain of $F_{(f), \vec{g}}$ has measure < 1 and (f) fails $F_{(f), \vec{g}}$ via an f_j such that the sequence $(f)_{\vec{c}}, f_j, \vec{g}$ is in $X_{\prec i, 1+k'+k}$.

For item (i), we use the second item of Proposition 3.17 and the fact that a countable product of sets of measure 1 will have measure 1. Item (ii) is a consequence of the second item of Proposition 3.17 and item (iii) is a consequence of Lemma 3.22.

We now consider $\langle (f), \vec{g} \rangle$ in the intersection of these three sets. Let $[f']$ be an extension of $[f]_{\prec i}(f)$. Assume that $\Lambda_{[f']}\big(\lambda g\{e\}_{|f'|}(\Lambda_{[f']}, \mu, \vec{h}, g, (f_a)_{\vec{c}}, \vec{g}, \vec{b})\big)$ terminates with a computation of ordinal rank at most α_i.

First assume that $\Lambda_{[f']}(F_{(f),\vec{g}}) = (f_{i'})$ for some $i' \prec i$. Then, for all $i'' \preceq i'$ and all j we have that

$$\{e\}_{[f']}(\Lambda_{[f']}, \mu, f_{i'',j}, (f)_{\vec{c}}, \vec{g}, \vec{b}) \downarrow$$

by a computation of ordinal rank below α_i. Since $(f)_{\vec{c}}, \vec{g} \in X_{\prec i, k'+k}$, we may apply the induction hypothesis, and see that $\{e\}_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \mu, f_{i'',j}, (f)_{\vec{c}}, \vec{g}, \vec{b}) \downarrow$ by the same computation. Then by a computation of ordinal rank not exceeding α_i:

$$\Lambda_{[f]_{\prec i}(f)}(\lambda g.\{e\}_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \mu, \vec{h}, g, (f)_{\vec{c}}, \vec{g}, \vec{b})) = (f_j)$$

This is the one spot where we need extra parameters from $[f]_{\prec i}$, in his case $f_{i'',j}$, when we formulate the properties of $\vec{g}$ used at step α_i. Since $f_{i'',j}$ may already be in $\vec{h}$, this is also the spot where we need Convention 3.28.

Secondly, suppose that the assumption from the previous paragraph is not the case. By the argument above, we then have the following:

$$\{e\}_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \mu, f_{i',j}, (f)_{\vec{c}}, \vec{g}, \vec{b}) \downarrow$$

for all $i' \prec i$ and $j \in \mathbb{N}$. There are two sub-cases to consider:

(i) If the domain of $F_{(f),\vec{g}}$ has measure 1, we get that $\Lambda_{[f]_{\prec i}(f)}(F_{(f),\vec{g}}) = (f)$ since (f) is the first single sequence in the double sequence $[f]_{\prec i}(f)$ that is sufficient for $F_{(f),\vec{g}}$. By the definition of $F_{(f),\vec{g}}$, this observation verifies the induction step in this case.

(ii) If the domain of $F_{(f),\vec{g}}$ has measure < 1, there is one f_j for which $F_{(f),\vec{g}}$ does not terminate. In light of item (i), we have that $(f)_{\vec{c}}, f_j, \vec{g}$ is in $X_{\prec i, 1+k'+k}$, and using the induction hypothesis negatively, we see that $\{e\}(\Lambda_{[f']}, \mu, \vec{h}, f_j, \vec{g}, \vec{b})$ does not terminate before α_i. Since this value is required for the $\Lambda_{[f']}$-computation in question to terminate at all, the latter cannot terminate at stage α_i or earlier.

We are now through all cases, i.e. the proof of Lemma 3.29 is finished.   □

We now let $[f]$ and each $X_{i,k}$ be as constructed above. For each $k \in \mathbb{N}$, we define the intersection $X_k = \bigcap_{i \in W} X_{i,k}$.

Lemma 3.30. *For each k and $\vec{g} \in X_k$ we have $\mathbf{m}(\{g \mid g, \vec{g} \in X_{k+1}\}) = 1$.*

Proof. It suffices to show that $\mathbf{m}(\{g \mid g, \vec{g} \in X_{i,k+1}\}) = 1$ for co-finally many $i \in W$, and it is the requirement in item (i) from the proof of Lemma 3.29 in the treatment of Λ-application that does the trick. Let $\vec{g} \in A_k$ and consider $\{e\}(\Lambda, \mu, \vec{g}) = \Lambda(\lambda g.\{e_1\}(\Lambda, \mu, g, \vec{g}))$ for any e_1 of suitable arity. When we treat this case stepping from $X_{\prec i,k}$ to $X_{i,k}$, the aforementioned item (i) restricts our attention to $\vec{g}$ additionally satisfying $\mathbf{m}(\{g \mid g, \vec{g} \in X_{\prec a,k+1}\}) = 1$. In the limit, this required property thus holds.   □

Let Λ_0 be any weak fan functional, and let $[f]$ be as constructed in the proof of Lemma 3.29. We define Λ_1 as follows:

$$\Lambda_1(F) = \begin{cases} \Lambda_{[f]}(F) & \text{if defined} \\ \Lambda_0(F) & \text{otherwise} \end{cases} \tag{3.4}$$

and prove our main theorem as follows.

Theorem 3.31. *If $f : \mathbb{N} \to \mathbb{N}$ is computable in $\Lambda_1 + \mu$, then it is computable in μ.*

Proof. We will prove the stronger claim (3.5) below by induction on the length of the computation. We need some notation as follows. Let e be a Kleene index, let $\vec{b}$ be a sequence from $\mathbb{N}$ and let $\vec{g}$ of length k be a sequence from $\bigcap_{i \in W} X_{i,k}$ such that $\omega_1^{\mathsf{CK}} = \omega_1^{\mathsf{CK},\vec{g}}$. By Proposition 3.23.(iv), the final restriction does not alter the measure of the set. Now consider the claim:

$$\{e\}(\Lambda_1, \mu, \vec{g}, \vec{b}) = a \to (\exists i \in W)(\{e\}_{[f]_i}(\Lambda_{[f]_i}, \mu, \vec{g}, \vec{b}) = a). \tag{3.5}$$

The theorem follows from the claim (3.5) and the total instances $\lambda c.\{e\}(\Lambda_\mu, c)$.

We now prove the claim (3.5) by induction on the ordinal rank of the computation $\{e\}(\Lambda_1, \mu, \vec{g}, \vec{b}) = a$. The proof is split into cases according to which Kleene scheme e represents, and all cases except those for application of μ or Λ_1 are trivial. We will consider the two cases (3.6) and (3.7). First, we consider

$$\{e\}(\Lambda_1, \mu, \vec{g}, \vec{b}) = \mu(\lambda c.\{e_1\}(\Lambda_1, \mu, \vec{g}, c, \vec{b})). \tag{3.6}$$

Then, by the induction hypothesis, we have the following termination property:

$$(\forall c \in \mathbb{N})(\exists i \in W)\big[\{e_1\}_{[f]_i}(\Lambda_{[f]_i}, \vec{g}, c, \vec{b})\!\downarrow\big].$$

Since ω_1^{CK} is Σ_1-admissible relative to $\vec{g}$ (see footnote 6), there is a bound on how far out in W we need to go, i.e. $(\exists i \in W)(\forall c \in \mathbb{N})[\{e_1\}a_{[f]_{\prec i}}(\Lambda_{[f]_{\prec i}}, \vec{g}, c, \vec{b})\!\downarrow]$, and $\{e\}_{[f]_i}\big(\Lambda_{[f_i]}, \mu, \vec{h}, \vec{g}, \vec{b}\big)\!\downarrow$ follows.

For the second case, consider the following (involving a slight abuse of notation):

$$\{e\}(\Lambda_1, \mu, \vec{g}, \vec{b}) = \Lambda_1(\lambda g.\{e_1\}(\Lambda_1, \mu, g, \vec{g}, \vec{b})). \tag{3.7}$$

Since this is a classically valid Kleene computation, we have that $\lambda g.\{e_1\}(\Lambda_1, \mu, g, \vec{g}, \vec{b})$ is total. By Lemma 3.30 and the induction hypothesis, for almost all g there is an $i_g \in W$ such that $\{e_1\}_{[f]_{i_g}}\big(\Lambda_{[f]_{i_g}}, \mu, g, \vec{g}, \vec{b}\big)\!\downarrow$. Now consider the sequence

$$i \mapsto \mathbf{m}\big(\{g \mid \{e_1\}_{[f]_i}(\Lambda_{[f]_i}, \mu, g, \vec{g}, \vec{b})\!\downarrow\}\big).$$

This sequence is increasing, computable in $\vec{g}$ and μ, and has limit 1, implying that

$$(\forall k)(\exists i_k \in W)\big(\mathbf{m}(\{g \mid \{e_1\}_{[f]_{i_k}}(\Lambda_{[f]_{i_k}}, \mu, g, \vec{g}, \vec{b})\!\downarrow\}) > 1 - 2^{-k}\big).$$

Hence, by the fact that $\omega_1^{\mathsf{CK}} = \omega_1^{\mathsf{CK},\vec{g}}$ and is Σ_1-admissible in $\vec{g}$, we see that there must be $i \in W$ such that

$$\mathbf{m}(\{g \mid \{e_1\}_{[f]_i}(\Lambda_{[f]_i}, \mu, g, \vec{g}, \vec{b})\!\downarrow\}) = 1.$$

In this light, our construction guarantees that (f_i) is sufficient for $\lambda g.\{e_1\}(\Lambda_1, \mu, g, \vec{g}, \vec{b})$. Unless some $(f_{i'})$ already does the job for $i' \prec i$, we may conclude that

$$\{e\}(\Lambda_1, \mu, \vec{g}, \vec{b}) = (f_i) = \{e\}_{[f]_i}(\Lambda_{[f]_i}, \mu, \vec{g}, \vec{b}).$$

This ends the induction step, and we are done. $\qquad\square$

4. Reverse Mathematics of the special fan functional

We show how Θ (and its generalisations) can reach the current outer edge of RM (and its higher-order generalisation).

First of all, in Section 4.1, we investigate the strength of the combination of Θ and S^2, which will be seen to reach the current *upper limit* of RM. Indeed, we have shown in [34] that the combination $S^2 + \Theta$ computes Gandy's *Superjump*, a functional intimately connected to Δ_2^1-CA_0. As a complementary result, we show in Section 4.1 that the system Π_1^1-$\mathsf{CA}_0^\omega + \mathsf{QF}\text{-}\mathsf{AC}^{2,1} + \mathsf{HBU}$ behaves as follows: (i) it implies Δ_2^1-CA_0, and (ii) it proves the same Π_3^1-sentences as Π_2^1-CA_0. To establish these results, we derive $[\Pi_2^1$-$\mathsf{CA}_0]^{\mathrm{st}}$ in $\mathsf{P}_0 + \Pi_1^1$-$\mathsf{TRANS} + \mathsf{STP}$.

Secondly, Θ, STP, and HBU express the compactness of Cantor space and the unit interval (in various forms). Since the compactness of *function spaces* is essential to the study of the gauge integral (see e.g. [24, 26]), it is a natural question how strong such compactness properties are. As a first step, we study in Section 4.2 the strength of such a compactness property inspired by STP. In particular, we formulate a generalisation of Theorems 2.19 and 4.1 to higher types suggested by [51, 52]. As a result, the compactness of function spaces seems quite strong from the point of view of RM.

4.1. **At the limit of Reverse Mathematics.** In this section, we derive $[\Pi_2^1$-$\mathsf{CA}_0]^{\mathrm{st}}$ in $\mathsf{P}_0 + \Pi_1^1$-$\mathsf{TRANS} + \mathsf{STP}$, which is a result similar to Theorem 2.19. We obtain interesting corollaries involving Δ_2^1-CA_0 and Π_2^1-CA_0. We first discuss some of the history of Π_2^1-CA_0 and related systems.

The system Π_2^1-CA_0 appears in the study of the Reverse Mathematics of topology by Mummert and Simpson ([27]), who identify this system as the 'current limit' of RM. The coding used by Mummert and Simpson is however not unproblematic, as discussed by Hunter ([12]). Furthermore, it is known that Π_2^1-CA_0 is equivalent to Σ_2^1-DC_0 and Σ_2^1-SEP_0 by [54, VII.6.9 and VII.6.14].

To the best of our knowledge, Π_2^1-CA_0 is also the current limit of *ordinal analysis*; according to Rathjen ([40]), the strength of Π_2^1-CA_0 *dwarfs* that of Π_1^1-CA_0. By the following theorem and its corollaries, STP and HBU are however all that is needed to step from the latter system to the former (in various guises).

Theorem 4.1. *The system* $\mathsf{P}_0 + \Pi_1^1$-$\mathsf{TRANS} + \mathsf{STP}$ *proves* $[\Pi_2^1$-$\mathsf{CA}_0]^{\mathrm{st}}$.

Proof. As noted in [54, VII.6.14], ACA_0 proves that Π_2^1-CA_0 is equivalent to Σ_2^1-SEP, where the latter is: For $\varphi_1, \varphi_2 \in \Sigma_2^1$ not involving the variable Z^1,

$$(\forall n^0)(\neg\varphi_1(n) \vee \neg\varphi_2(n)) \to (\exists Z^1)(\forall n^0)\big[\varphi_1(n) \to n \in Z \wedge \varphi_2(n) \to n \notin Z\big]. \quad (4.1)$$

We shall prove $[\Sigma_2^1$-$\mathsf{SEP}]^{\mathrm{st}}$ in $\mathsf{P}_0 + \Pi_1^1$-$\mathsf{TRANS} + \mathsf{STP}$. Since $\mathsf{P}_0 + \Pi_1^0$-TRANS proves the axioms of ACA_0 relative to 'st', we obtain $[\Pi_2^1$-$\mathsf{CA}_0]^{\mathrm{st}}$.

Let $\varphi_i(n)$ be short for the formula $(\exists g_i^1)(\forall h_i^1)(\exists x_i^0)(f_i(\overline{h_i}x_i, \overline{g_i}x_i, n) = 0)$ and fix standard f_i^1 for $i = 1, 2$. Then assume $\big[(\forall n^0)(\neg\varphi_1(n) \vee \neg\varphi_2(n))\big]^{\mathrm{st}}$, which is

$$(\forall^{\mathrm{st}} n^0)\big[(\forall^{\mathrm{st}} g_1^1)(\exists^{\mathrm{st}} h_1^1)(\forall^{\mathrm{st}} x_1^0)(f_1(\overline{h_1}x_1, \overline{g_1}x_1, n) \neq 0)$$
$$\vee (\forall^{\mathrm{st}} g_2^1)(\exists^{\mathrm{st}} h_2^1)(\forall^{\mathrm{st}} x_2^0)(f_2(\overline{h_2}x_2, \overline{g_2}x_2, n) \neq 0)\big].$$

Using $(\mu_1)^{\text{st}}$, which follows[7] from Π^1_1-TRANS, the previous formula implies that:

$$(\forall^{\text{st}} n^0, g^1_1, g^1_2)\big[\ (\forall^{\text{st}} x^0_1)(f_1\big(\ \overline{\mu_1(\lambda\sigma_1.f_1)}x_1, \overline{g_1}x_1, n\big) \neq 0) \tag{4.2}$$
$$\vee\ (\forall^{\text{st}} x^0_2)(f_2\big(\ \overline{\mu_1(\lambda\sigma_2.f_2)}x_2, \overline{g_2}x_2, n\big) \neq 0)\big],$$

where we suppressed parameters, as the 'full' notation of $\lambda\sigma_i.f_i$ is $\lambda\sigma^{0^*}_i.f_i(\sigma_i, \overline{g_i}x_i, n)$. Now fix nonstandard N^0 and apply Π^0_1-TRANS to (4.2) to obtain:

$$(\forall^{\text{st}} n^0, g^1_1, g^1_2)\big[\ (\forall x^0_1 \leq N)(f_1\big(\ \overline{\mu_1(\lambda\sigma_1.f_1)}x_1, \overline{g_1}x_1, n\big) \neq 0) \tag{4.3}$$
$$\vee\ (\forall x^0_2 \leq N)(f_2\big(\ \overline{\mu_1(\lambda\sigma_2.f_2)}x_2, \overline{g_2}x_2, n\big) \neq 0)\ \big].$$

Now let $A_i(n, g_i)$ be the (equivalent to quantifier-free) following formula

$$(\forall x^0_i \leq N)(f_i(\ \overline{\mu_1(\lambda\sigma_i.f_i)}x_i, \overline{g_i}x_i, n) \neq 0),$$

and let $A(n, g_1, g_2)$ be the formula $A_1(n, g_1) \vee A_2(n, g_2)$, i.e. the formula in square brackets in (4.3). By assumption, $(\forall^{\text{st}} n^0, g^1_1, g^1_2)A(n, g_1, g_2)$. Now consider:

$$(\forall^{\text{st}} v^{1^*}, x^{0^*})(\exists w^{1^*}, y^{0^*})(\forall g^1 \in v, n^0 \in x) \tag{4.4}$$
$$\big[g \in w \wedge n \in y \wedge (\forall h_1, h_2 \in w, m \in y)A(m, h_1, h_2)\big].$$

Note that (4.4) holds by taking $w = v$ and $y = x$. Applying I to (4.4), we obtain

$$(\exists w^{1^*}, y^{0^*})(\forall^{\text{st}} g^1, n^0)\big[g \in w \wedge n \in y \wedge (\forall h_1, h_2 \in w, m \in y)A(m, h_1, h_2)\big], \tag{4.5}$$

which -intuitively speaking- provides two sequences w, y (of nonstandard length) encompassing all standard functions and standard numbers *and* such that all of its elements satisfy A. In particular, one can view (4.5) as obtained by applying overspill to (4.3) while making sure all standard functions are in w.

Next, define the set Z^1_0 (actually a binary sequence) as follows: $n \in Z_0 \leftrightarrow (\exists g_1 \in w)\neg A_1(n, g)$, where w is the sequence from (4.5). Note that the right-hand side of the equivalence is actually '$(\exists i^0 < |w|)\neg A_1(n, w(i))$', i.e. Z_0 is definable in P_0.

Let Z^1 be a standard set such that $Z_0 \approx_1 Z$ as provided by STP. Furthermore, since μ_1 is standard, we have the following implications (for standard n):

$$(\exists^{\text{st}} g^1_1)(\forall^{\text{st}} h^1_1)(\exists^{\text{st}} x^0_1)(f_1(\overline{h_1}x_1, \overline{g_1}x_1, n) = 0)$$
$$\rightarrow (\exists^{\text{st}} g^1_1)(\exists^{\text{st}} x^0_1)(f_1(\overline{\mu_1(\lambda\sigma_1.f_1)}x_1, \overline{g_1}x_1, n) = 0)$$
$$\rightarrow (\exists g^1_1 \in w)(\exists x^0_1 \leq N)(f_1(\overline{\mu_1(\lambda\sigma_1.f_1)}x_1, \overline{g_1}x_1, n) = 0)$$
$$\rightarrow (\exists g^1_1 \in w)\neg A_1(n, g_1) \rightarrow n \in Z_0 \rightarrow n \in Z.$$

Now, since y from (4.5) contains all standard numbers, the second conjunct of (4.5) implies (by definition) that for standard m (by the definition of A):

$$(\forall h_1 \in w)A_1(m, h_1) \vee (\forall h_2 \in w)A_2(m, h_2). \tag{4.6}$$

---

[7]In the definition of Π^1_1-TRANS, bring outside the standard quantifiers and apply HAC$_{\text{int}}$. Introduce standard quantifiers in the antecedent using Π^0_1-TRANS to obtain $(\mu_1)^{\text{st}}$.

Similarly, consider the following series of implications (for standard n):

$$(\exists^{\text{st}} g_2^1)(\forall^{\text{st}} h_2^1)(\exists^{\text{st}} x_2^0)(f_2(\overline{h_2}x_2, \overline{g_2}x_2, n) = 0)$$
$$\to (\exists^{\text{st}} g_2^1)(\exists^{\text{st}} x_2^0)(f_2(\overline{\mu_1(\lambda\sigma_2.f_2)}x_2, \overline{g_2}x_2, n) = 0)$$
$$\to (\exists g_2^1 \in w)(\exists x_2^0 \le N)(f_2(\overline{\mu_1(\lambda\sigma_2.f_2)}x_2, \overline{g_2}x_2, n) = 0)$$
$$\to (\exists g_2^1 \in w)\neg A_2(n, g_2) \tag{4.7}$$
$$\to (\forall g_1^1 \in w)A_1(n, g_1) \tag{4.8}$$
$$\to n \notin Z_0 \to n \notin Z.$$

Note that (4.8) follows from (4.7) by (4.6). Thus, Z is as required for $[\Sigma_2^1\text{-SEP}]^{\text{st}}$. $\square$

The following corollary was proved in [33] by using the fact that no type two functional (hence including μ_1) can compute an instance of Θ. Hence, we observe that the computability-theoretic approach 'scales' better than our above approach via Nonstandard Analysis, but the latter may be called 'conceptually simpler'.

Corollary 4.2. *The system* $\mathsf{P}_0 + \Pi_1^1\text{-TRANS}$ *cannot prove* STP.

Proof. The system $\mathsf{E\text{-}PRA}^\omega + (\mu_1)$ is a Π_3^1-conservative extension of $\Pi_1^1\text{-CA}_0$ by [42, Theorem 2.2]. Furthermore, let φ be an arithmetical sentence (resp. not) provable in $\Pi_2^1\text{-CA}_0$ (resp. $\Pi_1^1\text{-CA}_0$). Suppose $\mathsf{P}_0 + \Pi_1^1\text{-TRANS} \vdash \mathsf{STP}$ and note that $\mathsf{P}_0 + \Pi_1^1\text{-TRANS} \vdash \varphi$ by the theorem (and the fact that $\varphi \leftrightarrow \varphi^{\text{st}}$ given $\Pi_1^0\text{-TRANS}$). Since $\Pi_1^1\text{-TRANS}$ is converted into (μ_1) by term extraction, we obtain $\mathsf{RCA}_0^\omega + (\mu_1) \vdash \varphi$, a contradiction with the aforementioned conservation result for (μ_1). $\square$

To be absolutely clear, we now discuss what does, and more importantly, *what does not* follow from Theorem 4.1.

Remark 4.3. First of all, one of the main consequences of the *Transfer* axiom of IST is the equivalence $\varphi \leftrightarrow \varphi^{\text{st}}$ (for any internal φ with standard parameters). In the absence of the *full* axiom of *Transfer*, as is the case for e.g. the system in Theorem 4.1, this equivalence may no longer hold. Hence, the system from Theorem 4.1 does *not* necessarily prove $\Pi_2^1\text{-CA}_0$. By contrast, the former system *does* prove the *arithmetical* consequences of $\Pi_2^1\text{-CA}_0$, thanks[8] to $\Pi_1^0\text{-TRANS}$.

Secondly, an interesting corollary of Theorem 2.19 is that $(\mu^2) + (\exists\Theta)\mathsf{SFF}(\Theta)$ implies ATR_0 over RCA_0^ω (see Theorem 2.11). To obtain this corollary, one observes that $\mathsf{ATR}_0^{\text{st}}$ implies (using $\Pi_1^0\text{-TRANS}$) the following normal form:

$$(\forall^{\text{st}} X^1, f^1)(\exists^{\text{st}} Y^1)\big[\mathsf{WO}(X) \to H_f(X, Y)\big]. \tag{4.9}$$

One then applies term extraction to $\mathsf{P}_0 + \Pi_1^0\text{-TRANS} + \mathsf{STP} \vdash$ (4.9); omitting the extracted term, one obtains that $[(\mu^2) + (\exists\Theta)\mathsf{SFF}(\Theta)] \to \mathsf{ATR}_0$ over RCA_0^ω. However, we can only obtain the latter implication *because* $\mathsf{ATR}_0^{\text{st}}$ implies an equivalent normal form, namely (4.9), which is *highly similar* to ATR_0 itself. The existence of such a 'highly similar' normal form (given a relatively weak system) is exceptional in that e.g. $[\mathsf{WKL}]^{\text{st}}$, $[\Sigma_1^1\text{-SEP}]^{\text{st}}$, and $[\Pi_2^1\text{-CA}_0]^{\text{st}}$ do not[9] have them, to the best of

---

[8]For internal and *arithmetical* φ with standard parameters, $\mathsf{P}_0 + \Pi_1^0\text{-TRANS} \vdash [\varphi \leftrightarrow \varphi^{\text{st}}]$.

[9]Let $\mathsf{WKL}_{\mathsf{ns}}$ be the statement that a *standard* and infinite binary tree has a *standard* path if the former contains sequences of arbitrary length. Then $\mathsf{P}_0 + \mathsf{WKL}^{\text{st}}$ (resp. $\mathsf{P}_0 + \mathsf{WKL}_{\mathsf{ns}}$) has the proof-theoretic strength of WKL_0 (resp. ACA_0), i.e. $\mathsf{WKL}^{\text{st}} \nrightarrow \mathsf{WKL}_{\mathsf{ns}}$ over P_0.

our knowledge. More generally, applying a proof interpretation (on which term extraction as in Theorem 2.3 is based) to the proof of a theorem, tends to completely warp the latter, i.e. $\mathsf{ATR}_0^{\mathrm{st}}$ is the exception, not the rule.

In light of Remark 4.3, it seems the system from Theorem 4.1 cannot prove $\Pi_2^1\text{-}\mathsf{CA}_0$; we now derive 'the next best thing' $\Delta_2^1\text{-}\mathsf{CA}_0$ from the result in Theorem 4.1.

Corollary 4.4. *The system* $\mathsf{RCA}_0^\omega + \mathsf{QF\text{-}AC}^{2,1} + (\mu_1) + \mathsf{HBU}$ *proves* $\Delta_2^1\text{-}\mathsf{CA}_0$.

Proof. Note that $\Pi_2^1\text{-}\mathsf{SEP} \leftrightarrow \Delta_2^1\text{-}\mathsf{CA}_0$ over ACA_0 by [54, VII.6.14], where $\Pi_2^1\text{-}\mathsf{SEP}$ is (4.1) for $\varphi_1, \varphi_2 \in \Pi_2^1$. By Theorem 4.1, $\mathsf{P}_0 + \Pi_1^1\text{-}\mathsf{TRANS} + \mathsf{STP}$ proves $[\Pi_2^1\text{-}\mathsf{SEP}]^{\mathrm{st}}$. The antecedent of the latter has the form $(\forall^{\mathrm{st}}n^0)(\exists^{\mathrm{st}}g^1)(\forall^{\mathrm{st}}h^1)\varphi^{\mathrm{st}}(n, g, h)$, where φ^{st} is arithmetical. Hence, the antecedent in $[\Pi_2^1\text{-}\mathsf{SEP}]^{\mathrm{st}}$ may be strengthened to

$$(\exists^{\mathrm{st}}\Phi^{0\to1^*})(\forall n^0)(\exists g^1 \in \Phi(n))(\forall h^1)\varphi(n, g, h) \tag{4.10}$$

using $\Pi_1^0\text{-}\mathsf{TRANS}$. On the other hand, the consequent of $[\Pi_2^1\text{-}\mathsf{SEP}]^{\mathrm{st}}$ has the form

$$(\exists^{\mathrm{st}}Z^1)(\forall^{\mathrm{st}}n^0)(\exists^{\mathrm{st}}g^1)\underline{(\forall^{\mathrm{st}}h^1)\psi^{\mathrm{st}}(n, g, h, Z)}, \tag{4.11}$$

where ψ^{st} is arithmetical. Now apply $\Pi_1^1\text{-}\mathsf{TRANS}$ (which readily follows from $(\exists^{\mathrm{st}}\mu_1)\mathsf{MUO}(\mu_1)$) to the underlined formula in (4.11). In the resulting formula, apply $\mathsf{HAC}_{\mathrm{int}}$ to obtain a standard functional $\Phi^{0\to1^*}$ such that:

$$(\exists^{\mathrm{st}}Z^1)(\forall^{\mathrm{st}}n^0)(\exists g^1 \in \Phi(n))\underline{(\forall h^1)\psi(n, g, h, Z)}, \tag{4.12}$$

Now apply $\Pi_1^0\text{-}\mathsf{TRANS}$ to the '$(\forall^{\mathrm{st}}n^0)$' quantifier in (4.12); note that $(\exists^{\mathrm{st}}\mu_1)\mathsf{MUO}(\mu_1)$ guarantees that the formula following the '$(\forall^{\mathrm{st}}n^0)$' quantifier is equivalent to a quantifier-free one. Thus, we obtain:

$$(\exists^{\mathrm{st}}\Psi^{0\to1^*}, Z^1)(\forall n^0)(\exists g^1 \in \Psi(n))(\forall h^1)\psi(n, g, h, Z) \tag{4.13}$$

using $(\exists^{\mathrm{st}}\mu_1)\mathsf{MUO}(\mu_1)$ and $\mathsf{HAC}_{\mathrm{int}}$. Now apply term extraction to

$$\mathsf{P}_0 + (\exists^{\mathrm{st}}\mu_1)\mathsf{MUO}(\mu_1) + \mathsf{STP} \vdash [(4.10) \to (4.13)]$$

and omit all terms. Finally note that $(\mu_1) + \mathsf{QF\text{-}AC}^{0,1}$ yields $\Phi^{0\to1^*}$ satisfying $(\forall n^0)(\exists g^1 \in \Phi(n))(\forall h^1)\varphi(n, g, h)$ from $(\forall n^0)(\exists g^1)\underline{(\forall h^1)\varphi(n, g, h)}$, as the underlined formula may be treated as quantifier-free. One thus obtains that $\mathsf{RCA}_0^\omega + \mathsf{QF\text{-}AC}^{0,1} + (\mu_1) + (\exists\Theta)\mathsf{SFF}(\Theta)$ proves $\Pi_2^1\text{-}\mathsf{SEP}$, and hence $\Delta_2^1\text{-}\mathsf{CA}_0$ as discussed above.   □

If one repeats the previous proof for $[\Sigma_2^1\text{-}\mathsf{SEP}]^{\mathrm{st}}$ (instead of $[\Pi_2^1\text{-}\mathsf{SEP}]^{\mathrm{st}}$), one will observe that *Transfer* for Π_2^1-formulas seems needed to treat the consequent of $[\Sigma_2^1\text{-}\mathsf{SEP}]^{\mathrm{st}}$ in the same way as in the previous proof. However, this instance of *Transfer* of course yields $\Pi_2^1\text{-}\mathsf{CA}_0$ after term extraction. In other words, the system from the corollary does not imply $\Pi_2^1\text{-}\mathsf{CA}_0$ *using the same proof*. We do obtain the following corollary where $\Pi_1^1\text{-}\mathsf{TR}_0$ is transfinite recursion for Π_1^1-formulas (see [54, VI.7.1]), i.e. ATR_θ from Section 4.9 for any $\theta \in \Pi_1^1$.

Corollary 4.5. *The system* $\mathsf{RCA}_0^\omega + (\mu_1) + (\exists\Theta)\mathsf{SFF}(\Theta)$ *proves* $\Pi_1^1\text{-}\mathsf{TR}_0$.

Proof. It is known that $\Pi_2^1\text{-}\mathsf{CA}_0$ implies $\Pi_1^1\text{-}\mathsf{TR}_0$ (see e.g. [54, VII.7.12]). By Theorem 4.1, $\mathsf{P}_0 + \Pi_1^1\text{-}\mathsf{TRANS} + \mathsf{STP}$ proves $[\Pi_1^1\text{-}\mathsf{TR}_0]^{\mathrm{st}}$. Similar to the second part of Remark 4.3, one observes that $\Pi_1^1\text{-}\mathsf{TR}_0^{\mathrm{st}}$ implies (using $\Pi_1^1\text{-}\mathsf{TRANS}$) the following normal form:

$$(\forall^{\mathrm{st}}X^1)(\exists^{\mathrm{st}}Y^1)\big[\mathsf{WO}(X) \to H_\theta(X, Y)\big], \tag{4.14}$$

for any fixed $\theta \in \Pi_1^1$. One then applies term extraction to $\mathsf{P}_0 + \Pi_1^1\text{-}\mathsf{TRANS} + \mathsf{STP} \vdash$ (4.14); omitting the extracted term, one obtains the corollary.   □

We now discuss some interesting proof-theoretic corollaries. Let $\mathsf{con}(S)$ be the Π_1^0-sentence expressing the consistency of S (see [54, II.8.2]).

Corollary 4.6. *The system* $\mathsf{RCA}_0^\omega + \mathsf{QF\text{-}AC}^{2,1} + (\mu_1) + \mathsf{HBU}$ *proves* $\mathsf{con}(\Pi_1^1\text{-}\mathsf{CA}_0)$; *the same holds for any* Π_3^1-*sentence provable in* Π_2^1-CA_0.

Proof. Since Π_2^1-$\mathsf{CA}_0 \vdash \mathsf{con}(\Pi_1^1$-$\mathsf{CA}_0)$, the system $\mathsf{P}_0 + \Pi_1^1$-$\mathsf{TRANS} + \mathsf{STP}$ proves $[\mathsf{con}(\Pi_1^1$-$\mathsf{CA}_0)]^{\mathrm{st}}$ and applying Π_1^0-TRANS yields $\mathsf{con}(\Pi_1^1$-$\mathsf{CA}_0)$. Hence, by Theorem 2.6, the stronger system $\mathsf{P}_0 + (\exists^{\mathrm{st}}\mu_1)\mathsf{MUO}(\mu_1) + (\exists^{\mathrm{st}}\Theta)\mathsf{SFF}(\Theta)$ proves $\mathsf{con}(\Pi_1^1$-$\mathsf{CA}_0)$. Applying term extraction as in Theorem 2.3, the corollary follows. For a Π_3^1-sentence $A \equiv (\forall X^1)(\exists Y^1)(\forall Z^1)\varphi(X, Y, Z)$, note that Π_1^1-TRANS yields $A^{\mathrm{st}} \leftrightarrow (\forall^{\mathrm{st}}X^1)(\exists^{\mathrm{st}}Y^1)(\forall Z^1)\varphi(X, Y, Z)$. Hence, if Π_2^1-$\mathsf{CA}_0 \vdash A$, the same proof as for $\mathsf{con}[\Pi_1^1$-$\mathsf{CA}_0]$ yields that $\mathsf{RCA}_0^\omega + (\mu_1) + (\exists\Theta^3)\mathsf{SFF}(\Theta)$ proves A. $\square$

This corollary is interesting as (μ_1) yields a conservative extension of Π_1^1-CA_0 (see [42, Theorem 2.2]), while HBU is acceptable in intuitionistic mathematics, and finitistically reducible (in the sense of yielding a conservative extension of WKL_0).

Corollary 4.7. *The systems* $\mathsf{P} + \Pi_1^1$-$\mathsf{TRANS} + \mathsf{STP}$ *and* $\mathsf{E\text{-}PA}^{\omega*} + (\mu_1) + \mathsf{QF\text{-}AC}^{2,1} + \mathsf{HBU}$ *prove the consistency of* Π_2^1-CA_0, *i.e.* $\mathsf{con}(\Pi_2^1$-$\mathsf{CA}_0)$.

Proof. Note that [54, VII.6.21] states Π_2^1-$\mathsf{CA}_0 \equiv_{\Pi_3^1} \Sigma_3^1$-$\mathsf{CA}_0$ and Π_2^1-$\mathsf{CA}_0 + \Sigma_3^1$-$\mathsf{IND} \vdash \mathsf{con}(\Sigma_3^1$-$\mathsf{CA}_0)$. Since Σ_3^1-$\mathsf{CA}_0 \rightarrow \Pi_2^1$-$\mathsf{CA}_0$ by [54, VII.6.6], the corollary follows. $\square$

Finally, by way of *mathematical* applications of Corollary 4.6, the *graph minor theorem* is a Π_1^1-sentence provable in Π_1^1-$\mathsf{CA}_0 + \mathsf{BI}$ ([8]); the latter system is derivable in Π_2^1-CA_0, yielding the following corollary.

Corollary 4.8. $\mathsf{RCA}_0^\omega + (\mu_1) + \mathsf{QF\text{-}AC}^{2,1} + \mathsf{HBU}$ *proves the graph minor theorem.*

4.2. **Generalisations to higher types.** In this section, we study compactness properties of function spaces. In particular, we study a generalisation of Theorems 2.19 and 4.1 to higher types inspired by [51, 52]. We first discuss the results in the latter and its relation to our results. We discuss the mathematical naturalness of compactness properties of function spaces in Remark 4.12.

First of all, recall that Theorem 2.19 was first proved in [33] by proving $[\Sigma_1^1$-$\mathsf{SEP}]^{\mathrm{st}}$ in $\mathsf{P}_0 + \Pi_1^0$-$\mathsf{TRANS} + \mathsf{STP}$, where Σ_1^1-SEP states that for any $\varphi_1, \varphi_2 \in \Sigma_1^1$:

$$(\forall n^0)(\neg\varphi_1(n) \vee \neg\varphi_2(n)) \rightarrow (\exists Z^1)(\forall n^0)\big(\varphi_1(n) \rightarrow n \in Z \wedge \varphi_2(n) \rightarrow n \notin Z\big).$$

The equivalence $\mathsf{ATR}_0 \leftrightarrow \Sigma_1^1$-$\mathsf{SEP}$ in [54, V.5.1] guarantees that $\mathsf{ATR}_0^{\mathrm{st}}$ is provable in $\mathsf{P}_0 + \Pi_1^0$-$\mathsf{TRANS} + \mathsf{STP}$. In this section, we study the higher type generalisation of $\mathsf{P}_0 + \Pi_1^0$-$\mathsf{TRANS} + \mathsf{STP} \vdash [\Sigma_1^1$-$\mathsf{SEP}]^{\mathrm{st}}$, inspired by results in [51, 52], sketched next.

Schweber discusses a higher-order generalisation of the RM of ATR_0 in [51, 52]. This generalisation consists in taking theorems from second-order arithmetic and 'bumping up all types with one' to obtain a theorem of third-order arithmetic. By way of example, compare Σ_1^1-SEP to the 'one level up' separation principle Σ_1^2-SEP (which is still provable in ZF) as follows.

Definition 4.9 (Σ_1^2-SEP). *For any* $\varphi_1, \varphi_2 \in \Sigma_1^2$, *we have that*

$$(\forall f^1)(\neg\varphi_1(f) \vee \neg\varphi_2(f)) \rightarrow (\exists Z^2)(\forall f^1)\big(\varphi_1(f) \rightarrow Z(f) = 1 \wedge \varphi_2(f) \rightarrow Z(f) = 0\big).$$

As noted by Schweber ([51]), Σ_1^2-SEP implies Δ_1^2-comprehension, and two determinacy axioms $\Sigma_1^{\mathbb{R}}$-DET and $\Delta_1^{\mathbb{R}}$-DET when combined with the axiom of choice as in SF($\mathbb{R}$). As noted by Hachtman in [9, 10], $\Sigma_1^{\mathbb{R}}$-DET is strictly stronger than Σ_4^0-DET, and Π_3^0-DET already goes beyond second-order arithmetic ([23, Cor. 1.3]).

As observed in [51, §1], many implications in the Reverse Mathematics of ATR$_0$ fail when the theorems are generalised from second-order to third-order arithmetic. It is thus a natural question if the implication $[\Pi_1^0\text{-TRANS} + \text{STP}] \to [\Sigma_1^1\text{-SEP}]^{\text{st}}$ generalises to third-order arithmetic. We answer this question positively as follows: The (obvious) generalisation of the system $\mathsf{P}_0 + \Pi_1^0\text{-TRANS} + \text{STP}$ to third-order arithmetic is $\mathsf{P}_0 + \text{SOT} + \text{STP}_2$ where the latter axioms are:

$$(\forall^{\text{st}} Y^2)\big[(\exists f^1)(Y(f) = 0) \to (\exists^{\text{st}} f^1)(Y(f) = 0)\big], \qquad \text{(SOT)}$$

$$(\forall Y^2 \leq_2 1)(\exists^{\text{st}} Z^2 \leq_2 1)(Z \approx_2 Y), \qquad \text{(STP}_2\text{)}$$

which are respectively Π_1^0-TRANS and STP with all types 'bumped up by one'. Recall that '$Z \approx_2 Y$' is $(\forall^{\text{st}} g^1)(Z(g) =_0 Y(g))$. We have the following theorem.

Theorem 4.10. *The system* $\mathsf{P}_0 + \text{SOT} + \text{STP}_2$ *proves* $[\Sigma_1^2\text{-SEP}]^{\text{st}}$.

Proof. Let $\varphi_i(f^1)$ be short for the formula $(\exists Y_i^2)(\forall f_i^1)(\psi_i^3(Y_i, f_i, f) = 0)$ and fix standard ψ_i^3 for $i = 1, 2$. Then assume $\big[(\forall f^1)(\neg\varphi_1(f) \vee \neg\varphi_2(f))\big]^{\text{st}}$, which is:

$$(\forall^{\text{st}} f^1)\big[(\forall^{\text{st}} Y_1^2)(\exists^{\text{st}} f_1^1)(\psi_1(Y_1, f_1, f) \neq 0) \vee (\forall^{\text{st}} Y_2^2)(\exists^{\text{st}} f_2^1)(\psi_2(Y_2, f_2, f) \neq 0)\big].$$

Now fix nonstandard u^{1^*} containing all standard sequences (which exists by *Idealisation* I) and note that we have that for all standard f^1, Y_1^2, Y_2^1:

$$(\exists f_1^1 \in u)(\psi_1(Y_1, f_1, f) \neq 0) \vee (\exists f_2^1 \in u)(\psi_2(Y_2, f_2, f) \neq 0) \qquad (4.15)$$

Let $A_i(f, Y_i)$ be the (equivalent to quantifier-free) formula $(\exists f_i^1 \in u)(\psi_1(Y_i, f_i, f) \neq 0)$ and let $A(f, Y_1, Y_2)$ be the formula $A_1(f, Y_1) \vee A_2(f, Y_2)$, i.e. the formula in (4.15). By assumption, $(\forall^{\text{st}} f^1, Y_1^2, Y_2^2)A(f, Y_1, Y_2)$. Now consider:

$$(\forall^{\text{st}} v^{2^*}, x^{1^*})(\exists w^{2^*}, y^{1^*})(\forall Y^2 \in v, f^1 \in x) \qquad (4.16)$$
$$\big[Y \in w \wedge f \in y \wedge (\forall Y_1, Y_2 \in w, f \in y)A(f, Y_1, Y_2)\big].$$

Note that (4.16) holds by taking $w =_{2^*} v$ and $y =_{1^*} x$. Applying I to (4.16) yields

$$(\exists w^{2^*}, y^{1^*})(\forall^{\text{st}} Y^2, f^1)\big[Y \in w \wedge f \in y \wedge (\forall Y_1, Y_2 \in w, f \in y)A(f, Y_1, Y_2)\big], \quad (4.17)$$

which -intuitively speaking- provides two sequences w, y (of nonstandard length) encompassing all standard functionals of type two and standard functions and such that all of its elements satisfy A. In particular, one can view (4.17) as obtained by applying overspill to (4.15) while making sure all standard functionals and functions are in w and y. Next, define the functional Z_0^2 as follows: $Z_0(f) = 0$ if $(\exists Y_1 \in w)\neg A_1(f, Y_1)$ and 1 otherwise, where w^{2^*} is the sequence from (4.17). Note that $(\exists Y_1 \in w)\neg A_1(f, Y_1)$ is actually '$(\exists i^0 < |w|)\neg A_1(f, w(i))$', i.e. Z_0^2 is definable in P_0.

Let Z^2 be a standard functional such that $Z_0 \approx_2 Z$ as provided by STP_2. Furthermore, SOT establishes the following implications (for standard f^1):

$$(\exists^{\text{st}} Y_1^2)(\forall^{\text{st}} f_1^1)(\psi_1(Y_1, f_1, f) = 0) \to (\exists^{\text{st}} Y_1^2)(\forall f_1^1)(\psi_1(Y_1, f_1, f) = 0)$$
$$\to (\exists^{\text{st}} Y_1^2)(\forall f_1^1 \in u)(\psi_1(Y_1, f_1, f) = 0)$$
$$\to (\exists Y_1^2 \in w)(\forall f_1^1 \in u)(\psi_1(Y_1, f_1, f) = 0)$$
$$\to (\exists Y_1^2 \in w)\neg A_1(f, Y_1) \to Z_0(f) = 0 \to Z(f) = 0.$$

Note that SOT is (only) necessary to establish the first implication. Now, since y from (4.17) contains all standard functions, the second conjunct of (4.17) implies (by definition) that for standard h^1 (by the definition of A):

$$(\forall Y_1^2 \in w) A_1(h, Y_1) \vee (\forall Y_2^2 \in w) A_2(h, Y_2). \tag{4.18}$$

Similarly, consider the following series of implications (for standard f^1):

$$(\exists^{\mathrm{st}} Y_2^2)(\forall^{\mathrm{st}} f_2^1)(\psi_2(Y_2, f_2, f) = 0) \rightarrow (\exists^{\mathrm{st}} Y_2^2)(\forall f_2^1)(\psi_1(Y_2, f_2, f) = 0)$$
$$\rightarrow (\exists^{\mathrm{st}} Y_2^2)(\forall f_2^1 \in u)(\psi_2(Y_2, f_2, f) = 0)$$
$$\rightarrow (\exists Y_2^2 \in w)(\forall f_2^1 \in u)(\psi_2(Y_2, f_2, f) = 0)$$
$$\rightarrow (\exists Y_2^2 \in w) \neg A_2(f, Y_2) \tag{4.19}$$
$$\rightarrow (\forall Y_1^2 \in w) A_1(f, Y_1) \tag{4.20}$$
$$\rightarrow Z_0(f) = 1 \rightarrow Z(f) = 1.$$

Note that SOT is (only) necessary to establish the first implication, while (4.20) follows from (4.19) by (4.18). Thus, we observe that Z^2 is as required for $\Sigma_2^1\text{-}\mathsf{SEP}$ relative to 'st', and we are done. $\qquad\square$

Note that $\mathsf{P}_0 + \mathsf{SOT}$ exists at the level of second-order arithmetic, while $\Sigma_1^2\text{-}\mathsf{SEP}$ goes beyond that. In other words, STP_2 yields a non-trivial step up in strength. The previous proof is readily generalised as follows: $[\Sigma_2^2\text{-}\mathsf{SEP}]^{\mathrm{st}}$ follows from STP_2 and *Transfer* for Σ_1^1-formulas. Finally, the axiom STP_2 has a normal form as follows.

Theorem 4.11. *In* P*,* STP_2 *is equivalent to*

$$(\forall^{\mathrm{st}} \Psi^{2 \rightarrow 1^*})(\exists^{\mathrm{st}} W^{2^*})(\forall Y^2 \leq_2 1)(\exists Z^2 \in W)(\forall f \in \Psi(Z))(Z(f) =_0 Y(f)). \tag{4.21}$$

Proof. Clearly, STP_2 implies (as standard sequences consist of standard elements):

$$(\forall^{\mathrm{st}} \Psi^{2 \rightarrow 1^*})(\forall Y^2 \leq_2 1)(\exists^{\mathrm{st}} Z^2 \leq_2 1)(\forall f \in \Psi(Z))(Z(f) =_0 Y(f)), \tag{4.22}$$

and the implication (4.22) $\rightarrow \mathsf{STP}_2$ is established as follows: Suppose $\neg\mathsf{STP}_2$, i.e. there is $Y_0^2 \leq_2 1$ such that $(\forall^{\mathrm{st}} Z^2 \leq_2 1)(\exists^{\mathrm{st}} f^1)(Z(f) \neq_0 Y(f))$. Applying $\mathsf{HAC}_{\mathrm{int}}$ to the latter, we obtain the negation of (4.22), and the latter is seen to be equivalent to STP_2. Finally, applying *Idealisation* I to (4.22), we obtain exactly (4.21). $\qquad\square$

The normal form (4.21) gives rise to the (non-unique) functional $\Sigma^{(2 \rightarrow 1^*) \rightarrow 2^*}$ defined by the following specification:

$$(\forall \Psi^{2 \rightarrow 1^*})(\forall Y^2 \leq_2 1)(\exists Z^2 \in \Sigma(\Psi))(\forall f \in \Psi(Z))(Z(f) =_0 Y(f)), \qquad (\mathsf{CFS}(\Sigma))$$

Intuitively, the open cover $\cup_{Y \in \{0,1\}^{\mathbb{N}^{\mathbb{N}}}} J_Y^{\Psi}$ has a finite sub-cover provided by $\Sigma(\Psi)$, where J_Y^{Ψ} is the neighbourhood of all $Z \in \{0,1\}^{\mathbb{N}^{\mathbb{N}}}$ which agree with Y on the finite sequence $\Psi(Y)$. In contrast[10] to special fan functionals, the functional Σ requires a non-trivial instance of the axiom of choice. The exact properties of Σ are beyond the scope of this paper and will be studied in a subsequent paper.

Finally, we discuss the mathematical naturalness of compactness properties of function spaces, and the associated gauge integrals.

---

[10]Define $\mathsf{SOT}(\xi) \equiv (\forall Y^2)\big[(\exists f^1)(Y(f) = 0) \rightarrow Y(\xi(Y)) = 0\big]$. Combining the results from [34, 50], any ξ^3 satisfying $\mathsf{SOT}(\xi)$ computes Θ via a term of Gödel's T, provable in $\mathsf{RCA}_0^\omega + (\exists^3)$. Note that $\exists^3$ introduced in Section 2.3 is a variation of such ξ.

Remark 4.12. The Feynman path integral is a central and fundamental object in physics, especially quantum mechanics. The Lebesgue integral does not provide an adequate formalisation for the path integral, but the latter *can* be formalised using the gauge integral ([24, 26]) over *function spaces*. As shown in [34, §3.3], compactness as in HBU is essential for the development of the gauge integral on the unit interval, and the compactness of function spaces is similarly essential for the formalisation of the Feynman path integral. However, as discussed in [25, §7], the compactness of function spaces can be treacherous waters. Hence, we only study STP_2 as above in this paper, and will establish the exact connection to the gauge integral in a later publication.

5. Conclusion

5.1. **Summary of results.** In this section, we provide a summary of the results in this paper and [33, 34]. Figure 1 below summarises these results concisely.

By way of a legend, in the right column are the linearly ordered 'Big Five' systems of RM, with above them full second-order arithmetic Z_2 and below them the system $WWKL_0 \equiv RCA_0 + WWKL$. In the middle column, we classify the functionals studied in this paper as follows: RCA_0^ω plus the existence of the pictured functional is (at least or exactly) at the level of the corresponding system on the right; (struck out) arrows denote (non) S1-S9-computability. In the left column, we classify the nonstandard axioms studied in this paper as follows: P_0 plus the pictured nonstandard axioms is (at least or exactly) at the level of the corresponding system on the right; (struck out) arrows denote (non)implication over P_0. Many questions regarding this diagram remain unanswered, as discussed in Section 5.2.

5.2. **Future research.** We discuss some open questions and future research.
  (i) The system $\Pi_2^1\text{-}CA_0 + \Pi_3^1\text{-}TI_0$ proves Δ_3^0-determinacy, while $\Pi_2^1\text{-}CA_0$ does not (see [21]). Hence, it is a natural question whether $P + STP + \Pi_1^1\text{-}TRANS$ proves transfinite induction as in $[\Pi_3^1\text{-}TI_0]^{st}$.
 (ii) What is the strength of nonstandard versions of *Hindman's theorem* ([11, §10.3.5])? The latter is strictly between ACA_0 and ATR_0.
(iii) What is the strength of nonstandard versions of DNR? Can these be derived from Λ and LMP?
 (iv) What is the strength of nonstandard versions of POS and 2-WWKL? What is their relation to Λ and LMP?
  (v) Combining $\exists^2$ or μ_1 with Θ results in a considerable jump in logical strength. Which functionals yield a similar jump in strength?
 (vi) Does the RM of WWKL give rise to interesting variations of Λ?
(vii) There are numerous theorems in classical analysis essentially of the form $(\forall x^2)(\exists y^{1/0})\Phi(x, y)$, and each of these defines a class of realisers $\zeta^{2\to 1/0}$ such that $(\forall x^2)\Phi(x, \zeta(x))$. A general investigation of the relative computational powers of such realisers, say modulo μ^2 or $\exists^2$, is warranted. We believe this study is intimately related to the RM study of the original theorems $(\forall x^2)(\exists y^{1/0})\Phi(x, y)$, and associated theorems from Nonstandard Analysis.

Furthermore, we have established a close link between Λ and the *Vitali covering lemma*, which we hope to develop further. Finally, the combination of Θ and the Suslin functional yields Gandy's *Superjump* ([34]), and we have additionally established that the former combination goes *far* beyond the latter functional. We hope
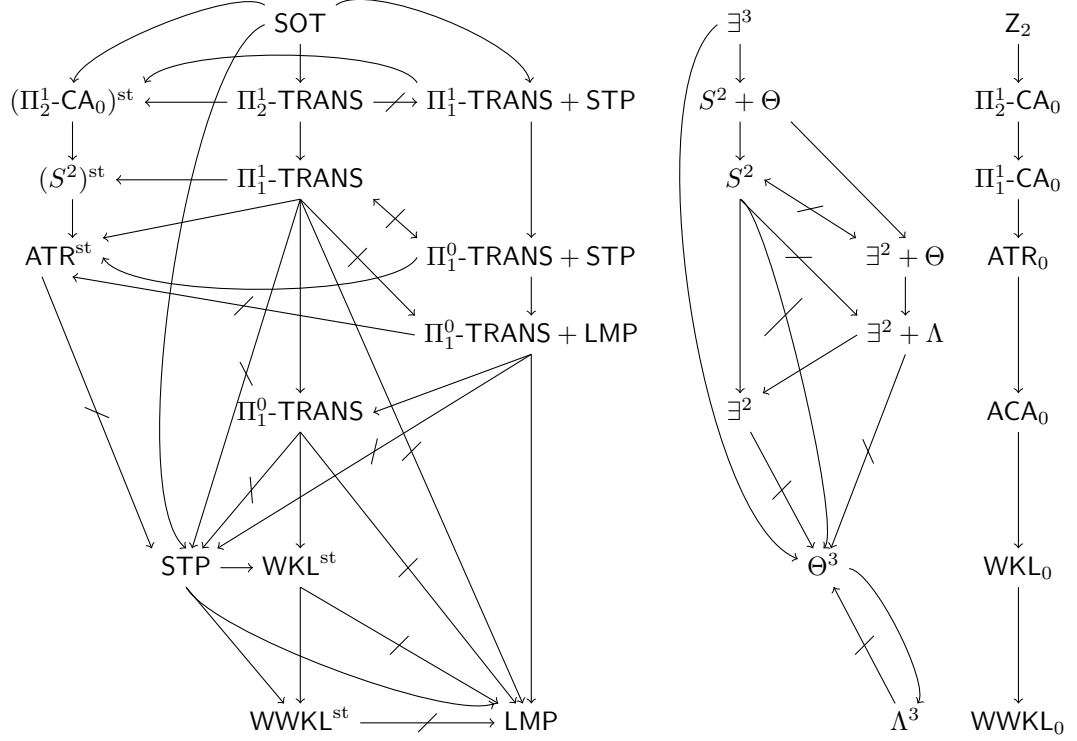
FIGURE 1. Summary of results

to establish the exact (logical and computational) strength of the aforementioned combination in the future.

APPENDIX A. SOME SYSTEMS OF NONSTANDARD ANALYSIS

In this section, we introduce Nelson's axiomatic approach to Nonstandard Analysis *internal set theory* ([31]), and it fragments based on Peano arithmetic from [4]. This background provides the definition for the systems P_0 and P used above.

A.1. **Internal set theory.** In Nelson's *syntactic* approach to Nonstandard Analysis ([31]), as opposed to Robinson's semantic one ([41]), a new predicate 'st(x)', read as 'x is standard' is added to the language of ZFC, the usual foundation of mathematics. The notations $(\forall^{st}x)$ and $(\exists^{st}y)$ are short for $(\forall x)(st(x) \to \dots)$ and $(\exists y)(st(y) \wedge \dots)$. A formula is called *internal* if it does not involve 'st', and *external* otherwise. The three external axioms *Idealisation*, *Standard Part*, and *Transfer* govern the new predicate 'st'; They are respectively defined[11] as:

(I) $(\forall^{st\ fin}x)(\exists y)(\forall z \in x)\varphi(z,y) \to (\exists y)(\forall^{st}x)\varphi(x,y)$, for any internal φ.
(S) $(\forall^{st}x)(\exists^{st}y)(\forall^{st}z)\big((z \in x \wedge \varphi(z)) \leftrightarrow z \in y\big)$, for any φ.

---

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

(T) $(\forall^{\mathrm{st}} t)\big[(\forall^{\mathrm{st}} x)\varphi(x,t) \to (\forall x)\varphi(x,t)\big]$, where $\varphi(x,t)$ is internal, and only has free variables t, x.

The system IST is just ZFC extended with the aforementioned external axioms; IST is a conservative extension of ZFC for the internal language, as proved in [31].

Clearly, the extension from ZFC to IST can also be done for *subsystems* of the former. Such extensions are studied in [4] for the classical and constructive formalisations of arithmetic, i.e. *Peano arithmetic* and *Heyting* arithmetic. In particular, the systems studied in [4] are E-HA$^\omega$ and E-PA$^\omega$, respectively *Heyting and Peano arithmetic in all finite types and the axiom of extensionality*. We refer to [16, §3.3] for the exact definitions of the (mainstream in mathematical logic) systems E-HA$^\omega$ and E-PA$^\omega$. We introduce in Section A.2 the system P, the (conservative) extension of E-PA$^\omega$ with fragments of the external axioms of IST.

Finally, E-PA$^{\omega*}$ is the definitional extensions of E-PA$^\omega$ with types for finite sequences, as in [4, §2]. For the former system, we require some notation.

Notation A.1 (Finite sequences)**.** The systems E-PA$^{\omega*}$ and E-HA$^{\omega*}$ have a dedicated type for 'finite sequences of objects of type ρ', namely ρ^*. Since the usual coding of pairs of numbers goes through in both, 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 $s^{\rho^*} = \langle s_0^\rho, s_1^\rho, \ldots, s_{n-1}^\rho\rangle$, 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 $(s * t)(j) = t(j - |s|)$ for $|s| \le j < |s| + |t|$. For a sequence s^{ρ^*}, we define $\bar{s}N := \langle s(0), s(1), \ldots, s(N)\rangle$ for $N^0 < |s|$. For a sequence $\alpha^{0\to\rho}$, we also write $\bar{\alpha}N = \langle \alpha(0), \alpha(1), \ldots, \alpha(N)\rangle$ 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 $\underline{x}, \underline{y}, \underline{t}, \ldots$ as short for tuples $x_0^{\sigma_0}, \ldots x_k^{\sigma_k}$ of possibly different type σ_i.

Remark A.2 (Notation)**.** The system E-PA$^{\omega*}$ includes equality between natural numbers '$=_0$' as a primitive. Equality '$=_\tau$' and inequality $\le_\tau$ for x^τ, y^τ is:

$$[x =_\tau y] \equiv (\forall z_1^{\tau_1} \ldots z_k^{\tau_k})[xz_1 \ldots z_k =_0 yz_1 \ldots z_k], \tag{A.1}$$

$$[x \le_\tau y] \equiv (\forall z_1^{\tau_1} \ldots z_k^{\tau_k})[xz_1 \ldots z_k \le_0 yz_1 \ldots z_k], \tag{A.2}$$

if the type τ is composed as $\tau \equiv (\tau_1 \to \ldots \to \tau_k \to 0)$. In the spirit of Nonstandard Analysis, we define 'approximate equality $\approx_\tau$' as follows (with the type τ as above):

$$[x \approx_\tau y] \equiv (\forall^{\mathrm{st}} z_1^{\tau_1} \ldots z_k^{\tau_k})[xz_1 \ldots z_k =_0 yz_1 \ldots z_k] \tag{A.3}$$

All the above systems include the *axiom of extensionality* for all $\varphi^{\rho\to\tau}$ as follows:

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

However, as noted in [4, p. 1973], the so-called axiom of *standard* extensionality $(\mathsf{E})^{\mathrm{st}}$ is problematic and cannot be included in P or P_0.

A.2. **The classical systems P and P_0.** We first introduce the system P, a conservative extension of $E\text{-}PA^\omega$ with fragments of Nelson's IST.

To this end, we first introduce the base system $E\text{-}PA^{\omega*}_{st}$. We use the same definition as [4, Def. 6.1], where $E\text{-}PA^{\omega*}$ is the definitional extension of $E\text{-}PA^\omega$ with types for finite sequences as in [4, §2]. The set $\mathcal{T}^*$ is defined as the collection of all the constants in the language of $E\text{-}PA^{\omega*}$.

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

(1) The schema[12] $st(x) \wedge x = y \to st(y)$,
(2) The schema providing for each closed term $t \in \mathcal{T}^*$ the axiom $st(t)$.
(3) The schema $st(f) \wedge st(x) \to st(f(x))$.

The external induction axiom IA^{st} states that for any (possibly external) Φ:

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

Secondly, we introduce some essential fragments of IST studied in [4].
Definition A.4.

(1) HAC_{int}: For any internal formula φ, we have

$$(\forall^{st} x^\rho)(\exists^{st} y^\tau)\varphi(x,y) \to \left(\exists^{st} F^{\rho \to \tau^*}\right)(\forall^{st} x^\rho)(\exists y^\tau \in F(x))\varphi(x,y), \qquad (A.4)$$

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

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

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

Note that I and HAC_{int} are fragments of Nelson's axioms *Idealisation* and *Standard part*. By definition, F in (A.4) only provides a *finite sequence* of witnesses to $(\exists^{st} y)$, explaining its name *Herbrandized Axiom of Choice*.

The system P is connected to $E\text{-}PA^\omega$ by Theorem 2.3 which expresses that we may obtain effective results as in (2.5) from any theorem of Nonstandard Analysis which has the same form as in (2.4). The scope of this theorem includes the Big Five systems of Reverse Mathematics ([46]), the Reverse Mathematics zoo ([49]), and both classical and higher-order computability theory ([44,47]).

We now introduce the system P_0, a conservative extension of RCA^ω_0 with fragments of Nelson's IST. Recall that the system $RCA^\omega_0 \equiv E\text{-}PRA^\omega + QF\text{-}AC^{1,0}$ is Kohlenbach's *base theory of higher-order Reverse Mathematics* as introduced in [17, §2]. The system $E\text{-}PRA^{\omega*}$ is an obvious definitional extensional as in Remark A.1. Recall that we permit ourselves a slight abuse of notation by also referring to $E\text{-}PRA^{\omega*} + QF\text{-}AC^{1,0}$ as RCA^ω_0.

Definition A.5. The system P_0 is $E\text{-}PRA^{\omega*} + QF\text{-}AC^{1,0} + \mathcal{T}^*_{st} + I + HAC_{int}$.

Finally, the system P_0 is connected to RCA^ω_0 by Corollary 2.4.

---

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

for the Promotion of Science. The authors express their gratitude towards these institutions. The authors thank Tom Powell, Martin Hyland, Ulrich Kohlenbach, and Anil Nerode for their valuable advice. The referee and editor were also instrumental in greatly improving this paper.

References

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

[2] 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.

[3] Vasco Brattka and Guido Gherardi, *Effective choice and boundedness principles in computable analysis*, Bull. Symbolic Logic **17** (2011), no. 1, 73–117.

[4] Benno van den Berg, Eyvind Briseid, and Pavol Safarik, *A functional interpretation for nonstandard arithmetic*, Ann. Pure Appl. Logic **163** (2012), no. 12, 1962–1994.

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

[6] Pierre Cousin, *Sur les fonctions de n variables complexes*, Acta Math. **19** (1895), no. 1, 1–61.

[7] Damir D. Dzhafarov, *Reverse Mathematics Zoo*. http://rmzoo.uconn.edu/.

[8] Harvey Friedman, Neil Robertson, and Paul Seymour, *The metamathematics of the graph minor theorem*, Contemp. Math., vol. 65, Amer. Math. Soc., 1987, pp. 229–261.

[9] Sherwood Julius Hachtman, *Calibrating Determinacy Strength in Borel Hierarchies*, ProQuest LLC, Ann Arbor, MI, 2015. Thesis (Ph.D.)–University of California, Los Angeles.

[10] ——, *Determinacy In Third Order Arithmetic* (2017), 12. http://homepages.math.uic.edu/~shac/reverse.pdf.

[11] Denis R. Hirschfeldt, *Slicing the truth*, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, vol. 28, World Scientific Publishing, 2015.

[12] James Hunter, *Higher-order reverse topology*, ProQuest LLC, Ann Arbor, MI, 2008. Thesis (Ph.D.)–The University of Wisconsin - Madison.

[13] 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.

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

[15] Stephen C. Kleene, *Recursive functionals and Quantifiers of finite types I*, Transactions of the AMS **91** (1959), 1–52.

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

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

[18] ——, *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).

[19] Ernst Lindelöf, *Sur Quelques Points De La Théorie Des Ensembles*, Comptes Rendus (1903), 697–700.

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

[21] MedYahya Ould MedSalem and Kazuyuki Tanaka, Δ_3^0-*determinacy, comprehension and induction*, J. Symbolic Logic **72** (2007), no. 2, 452–462.

[22] Fyodor A. Medvedev, *Scenes from the history of real functions*, Science Networks. Historical Studies, vol. 7, Birkhäuser Verlag, Basel, 1991.

[23] Antonio Montalbán and Richard A. Shore, *The limits of determinacy in second-order arithmetic*, Proc. Lond. Math. Soc. (3) **104** (2012), no. 2, 223–252.

[24] Pat Muldowney, *A general theory of integration in function spaces, including Wiener and Feynman integration*, Pitman Research Notes in Mathematics Series, vol. 153, Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1987.

[25] ——, *The infinite dimensional Henstock integral and problems of Black-Scholes expectation*, J. Appl. Anal. **8** (2002), no. 1, 1–21.

[26] ——, *A modern theory of random variation*, John Wiley & Sons, 2012. With applications in stochastic calculus, financial mathematics, and Feynman integration.

[27] Carl Mummert and Stephen G. Simpson, *Reverse mathematics and Π_2^1 comprehension*, Bull. Symbolic Logic **11** (2005), no. 4, 526–533.

[28] E. Nathanson, *Path integration with non-positive distributions and applications to the Schrödinger equation*, PhD (Doctor of Philosophy) thesis, University of Iowa, `https://doi.org/10.17077/etd.k483ok3i`. (2014).

[29] E. Nathanson and P. Jørgensen, *A global solution to the Schrödinger equation: From Henstock to Feynman*, Journal of Mathematical Physics **56** (2015).

[30] _____, *Trotter's limit formula for the Schrödinger equation with singular potential*, Journal of Mathematical Physics **58** (2017).

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

[32] Dag Normann, *Recursion on the countable functionals*, LNM 811, vol. 811, Springer, 1980.

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

[34] _____, *On the mathematical and foundational significance of the uncountable*, to appear in Journal of Mathematical Logic, arXiv: `https://arxiv.org/abs/1711.08939` (2018).

[35] _____, *Pincherle's theorem in Reverse Mathematics and computability theory*, Submitted, arXiv: `https://arxiv.org/abs/1808.09783` (2018).

[36] _____, *Representations in measure theory*, arXiv: `https://arxiv.org/abs/1902.02756` (2019).

[37] Salvatore Pincherle, *Notice sur les travaux*, Acta Math. **46** (1925), no. 3-4, 341–362.

[38] _____, *Sopra alcuni sviluppi in serie per funzioni analitiche (1882)*, Opere Scelte, I, Roma (1954), 64–91.

[39] W. N. Polyzou and Ekaterina Nathanson, *Scattering using real-time path integrals*, arXiv: `https://arxiv.org/abs/1712.00046` (2017).

[40] Michael Rathjen, *The art of ordinal analysis*, International Congress of Mathematicians. Vol. II, Eur. Math. Soc., Zürich, 2006.

[41] Abraham Robinson, *Non-standard analysis*, North-Holland, Amsterdam, 1966.

[42] Nobuyuki Sakamoto and Takeshi Yamazaki, *Uniform versions of some axioms of second order arithmetic*, MLQ Math. Log. Q. **50** (2004), no. 6, 587–593.

[43] Gerald E. Sacks, *Higher recursion theory*, Perspectives in Mathematical Logic, Springer, 1990.

[44] Sam Sanders, *The Gandy-Hyland functional and a computational aspect of nonstandard analysis*, Computability **7** (2018), no. 1, 7–43.

[45] _____, *Some nonstandard equivalences in reverse mathematics*, Sailing routes in the world of computation, Lecture Notes in Comput. Sci., vol. 10936, Springer, Cham, 2018, pp. 365–375.

[46] _____, *The unreasonable effectiveness of Nonstandard Analysis*, Submitted, Available from: `http://arxiv.org/abs/1508.07434` (2015).

[47] _____, *From Nonstandard Analysis to various flavours of computability theory*, To appear in Proceedings of TAMC17, Lecture Notes in Computer Science, Springer, Available from: `https://arxiv.org/abs/1609.01919` (2017).

[48] _____, *To be or not to be constructive*, Indagationes Mathematicae and the volume *L.E.J. Brouwer, fifty years later*, arXiv: `https://arxiv.org/abs/1704.00462` (2017), pp. 69.

[49] _____, *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).

[50] _____, *Higher-order computability and metastability*, Proceedings of LFCS18, LNCS 10703, Springer (2018).

[51] Noah Schweber, *Transfinite recursion in higher reverse mathematics*, J. Symb. Log. **80** (2015), no. 3, 940–969. `https://arxiv.org/abs/1310.5792`.

[52] _____, *Interactions between computability theory and set theory*, ProQuest LLC, Ann Arbor, MI, 2016. Thesis (Ph.D.)–University of California, Berkeley.

[53] Stephen G. Simpson (ed.), *Reverse mathematics 2001*, Lecture Notes in Logic, vol. 21, ASL, La Jolla, CA, 2005.

[54] _____, *Subsystems of second order arithmetic*, 2nd ed., Perspectives in Logic, CUP, 2009.

[55] Stephen G. Simpson and Keita Yokoyama, *A nonstandard counterpart of* WWKL, Notre Dame J. Form. Log. **52** (2011), no. 3, 229–243.

[56] John Stillwell, *Reverse mathematics, proofs from the inside out*, Princeton Univ. Press, 2018.

[57] Terence Tao, *Compactness and Compactification* (Timothy Gowers, ed.), The Princeton Companion to Mathematics, Princeton University Press, 2008.

[58] Anne Sjerp Troelstra, *Metamathematical investigation of intuitionistic arithmetic and analysis*, Springer Berlin, 1973. Lecture Notes in Mathematics, Vol. 344.

[59] Alan Turing, *On computable numbers, with an application to the Entscheidungs-problem*, Proceedings of the London Mathematical Society **42** (1936), 230-265.

[60] Guiseppe Vitali, *Sui gruppi di punti e sulle funzioni di variabili reali.*, Atti della Accademia delle Scienze di Torino, vol XLIII **4** (1907), 229–247.

[61] Karl Weierstrass, *Ausgewählte Kapitel aus der Funktionenlehre*, Teubner-Archiv zur Mathematik, vol. 9, BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1988.