

# ON THE MATHEMATICAL AND FOUNDATIONAL SIGNIFICANCE OF THE UNCOUNTABLE

DAG NORMANN AND SAM SANDERS

ABSTRACT. For various historical, practical, and foundational reasons, large parts of mathematics are studied *indirectly* via *countable* approximations, also called *codes*. It is a natural question whether this indirect study of codes is faithful to the original development in mathematics, or whether approximations somehow distort the latter. Another natural question is which parts of basic mathematics can(not) be studied via these representations. In this paper, we formulate new answers to these old questions. Our answers stem both from mathematics itself (via the study of the *gauge integral*) and its foundations (via Hilbert-Bernays' *Grundlagen der Mathematik* and its spiritual successor *Reverse Mathematics*). We identify a number of basic theorems from (uncountable) mathematics for which the logical and computational properties are *completely* (and even *maximally*) distorted upon introducing countable approximations. In a nutshell, while countable approximations are interesting and important, even extremely basic 'uncountable' mathematics is infinitely more complicated than the 'countable picture' involving codes suggests.

## 1. INTRODUCTION

**1.1. Motivation and goal.** Lest there be any doubt about our motivation, we start with a clear caveat: we firmly believe countable approximations are interesting and important; *however*, even basic uncountable mathematics is, as it turns out, infinitely more complicated than the associated 'countable picture' suggests. The goal of this paper is to communicate some of this fascinating (and surprising) complexity of basic uncountable mathematics to an audience of mathematicians.

By way of background, it is a commonplace that large parts of mathematics are studied *indirectly* via *countable approximations*. A prominent example is the program *Reverse Mathematics*, as founded by Friedman and developed extensively by Simpson (See Section 2.1). The underlying framework is *second-order arithmetic*  $Z_2$ , in which only natural numbers and sets thereof are available. Thus,  $Z_2$  is *by design* restricted to countable objects, but the latter can be used to *indirectly* represent e.g. continuous functions on  $\mathbb{R}$ , separable spaces, and even some topologies. Second-order arithmetic interprets a strong<sup>1</sup> fragment of Zermelo-Fraenkel set theory (the usual foundations of mathematics), and due to its ability to formalise analysis,  $Z_2$  is also called "analysis" ([53, p. 291]). Finally, countable representations are also called *codes*, and the use of codes is generally deemed to be *unproblematic*.

By contrast, the goal of this paper is to show that the use of codes *fundamentally* distorts mathematics. To this end, we provide a list of basic theorems (See Section 1.3) for which the logical and computational properties are *completely* (and in

---

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF OSLO, P.O. BOX 1053, BLINDERN N-0316  
CENTER FOR ADVANCED STUDIES AND MCMP, LMU MUNICH, GERMANY

*E-mail addresses:* `dnormann@math.uio.no`, `sasander@me.com`.

<sup>1</sup>It is known that  $Z_2$  interprets the fragment of Zermelo-Fraenkel set theory *without* the axiom of choice and *without* the axiom of powerset ([36]).

a precise sense *maximally*) distorted upon introducing countable approximations. Some of the listed theorems are well-established: the oldest (and most relevant) theorem is the *Cousin lemma* from [13, p. 22], dating back about 125 years, while the *Lindelöf lemma* ([34, p. 698]) dates back about 115 years. Some of the listed theorems are of great conceptual importance as they pertain to the *gauge integral* ([3]), which provides a formalisation of the *Feynman path integral* ([11, 38]).

Intuitively speaking, we show that the basic theorems from Section 1.3 live at the outer edge of second-order arithmetic, while their ‘countable substitutes’ (if any) live in very weak fragments of  $\mathbf{Z}_2$ . Besides being a *huge* leap in logical strength, our results also call into question the ‘Big Five’ classification from Reverse Mathematics (See Section 2.1). Nonetheless, we obtain in Section 3.3 a number of *Reverse Mathematics style* equivalences involving the Cousin lemma and *basic* properties of the *gauge integral*, like its uniqueness and its extension of the Riemann integral.

On a conceptual note, Reverse Mathematics is intimately connected to *classical* computability theory; similarly, our results have an (almost) equivalent reformulation in *higher-order* computability theory, and are even (often) obtained via the latter. Furthermore, in light of this correspondence, we investigate in Section 4 the strength of basic theorems like the Cousin and Lindelöf lemmas when combined with fundamental objects from computability theory. This study yields surprising results reaching all the way up to Gandy’s *superjump* ([23]), a ‘higher-order’ version of Turing’s *Halting problem* ([58]), the prototypical non-computable object.

Next, we discuss the motivations for this paper (stemming both from mathematics and its foundations) in detail in Section 1.2.

**1.2. Foundational and mathematical motivations.** We discuss the motivations for this paper. Items (i) and (ii) motivate the study of mathematics beyond second-order arithmetic, while a notable consequence is provided by item (iii).

- (i) The *gauge integral* is a generalisation of the Lebesgue and (improper) Riemann integral, and formalises Feynman’s path integral (See Section 1.2.1). Second-order arithmetic cannot accommodate (basic) gauge integration.
- (ii) The foundational studies of mathematics led by Hilbert take place in a logical framework *richer* than second-order arithmetic (See Section 1.2.2). It is natural to ask if anything is lost by restricting to the latter.
- (iii) The *compatibility problem* for Nelson’s *predicative arithmetic* ([41]) was solved in the negative ([12]). We solve the compatibility problem for *Weyl-Feferman predicative mathematics* in the negative (See Section 1.2.3).

As an example of how items (i) and (ii) are (intimately) related: the *uniqueness* of the gauge integral requires (Heine-Borel) compactness for *arbitrary* open covers. The latter compactness cannot be formulated in second-order arithmetic, and will be seen to have *completely* different logical and computational properties compared to the ‘countable’ substitute, i.e. (Heine-Borel) compactness for *countable* covers.

**1.2.1. Mathematical motivations.** In this section, we discuss the mathematical motivations for this paper, provided by the study of the *gauge integral*. As will become clear, the study of the latter goes beyond second-order arithmetic.

First of all, the gauge integral (aka *Henstock-Kurzweil* integral) was introduced around 1912 by Denjoy (in a different form) and constitutes a simultaneous generalisation of the Lebesgue and (improper) Riemann integral. The gauge integral

provides a formal framework for the *Feynman path integral* ([11, 38]), i.e. gauge integrals are also highly relevant in (the foundations of) physics. As expected, the gauge integral can handle discontinuous functions, which were around at the time already: Dirichlet discusses the characteristic function of the rationals around 1829 in [14], while Riemann defines a function with countably many discontinuities via a series in his *Habilitationsschrift* (See [31, p. 115]).

Secondly, since Lebesgue integration is studied in Reverse Mathematics (See [55, X.1]), it is a natural next step to study the gauge integral. However, even the definition of the latter goes beyond second-order arithmetic, as shown in Section 3.3; the same holds for general discontinuous functions which are needed to prove basic additivity results for the gauge integral. Furthermore, by Theorem 3.19, the *uniqueness* of the gauge integral requires the *Cousin lemma* ([13, p. 22]), which deals with *uncountable covers*, i.e. also beyond second-order arithmetic.

In conclusion, the gauge integral seems to require a logical framework *richer* than second-order arithmetic. As suggested above, this study will yield surprising results: the Cousin lemma expresses compactness for *arbitrary* open covers; this lemma turns out to have *completely* different logical and computational properties compared to the ‘countable’ substitute, i.e. compactness for *countable* covers.

1.2.2. *Foundational motivations.* We show that the foundational studies of mathematics led by Hilbert took place in a framework *richer* than second-order arithmetic. First of all, in his 1917-1933 lectures on the foundations of mathematics ([27]), Hilbert used a logical system involving third-order *Funktionfunktionen*, i.e. beyond second-order arithmetic. Ackermann’s 1924 dissertation (supervised by Hilbert) starts with an overview of *Hilbertsche Beweistheorie*, i.e. Hilbertian proof theory, which explicitly includes third-order parameters and the ‘epsilon’ operator.

Secondly, Hilbert and Bernays introduce<sup>2</sup> the formal system  $H$  in [29, Supplement IV], and use it to formalise parts of mathematics. The ‘epsilon’ operator is again at the core of this formalism; this operator is used by Hilbert and Bernays to define a certain object  $\xi$  in [29, p. 495]. Now,  $\xi$  maps functions to functions and is hence a third-order object, not available in second-order arithmetic in particular.

Thirdly, Simpson<sup>3</sup> positions *Reverse Mathematics* (See Section 2.1) in [55, p. 6] as a continuation of Hilbert-Bernays’ research, namely as follows:

The development of a portion of ordinary mathematics within [second-order arithmetic]  $Z_2$  is outlined in Supplement IV of Hilbert/Bernays [115]. The present book may be regarded as a continuation of the research begun by Hilbert and Bernays.

In conclusion, the foundational studies of Hilbert-Bernays-Ackermann take place in a logical framework *richer* than second-order arithmetic, and it is a natural *foundational* question if anything is lost by restricting to the latter.

1.2.3. *Foundational consequences.* We discuss the *compatibility problem* for predicative mathematics *à la* Weyl-Fefermann. As it turns out, our results solve this problem in the negative, providing another motivation for this paper.

<sup>2</sup>All other systems in [29, Suppl. IV] are either a variation of  $H$  or more limited than  $H$ .

<sup>3</sup>Note that the reference ‘[115]’ in the quote is [28, 29] in Simpson’s list of references in [55].

Russell famously identified an inconsistency in early set theory, known as *Russell's paradox*, based on the ‘set of all sets’ ([59]). According to Russell, the source of this paradox was *circular reasoning*: in defining the ‘set of all sets’, one quantifies over all sets, including the one that is being defined. To avoid such problems, Russell suggested banning any *impredicative definition*, i.e. a definition in which one quantifies over the object being defined. The textbook example of an impredicative definition is the supremum of a bounded set of reals, defined as the *least* upper bound of that set. Weyl, a student of Hilbert, initiated the development of *predicative mathematics* ([61]), i.e. avoiding impredicative definitions, which Feferman continued ([17–19]). Finally, the fourth ‘Big Five’ system of Reverse Mathematics is considered the ‘upper limit’ of predicative mathematics (See [55, §I.12]).

In an (similar but much more strict) effort to develop mathematics based on a predicative notion of *number*, Nelson introduced *predicative arithmetic* ([41]). Unfortunately, predicative arithmetic suffers from the *compatibility problem*: If two theorems  $A, B$  are both acceptable from the point of view of predicative arithmetic, it is possible that  $A \wedge B$  is not ([12]). In this light, the development of predicative arithmetic seems somewhat arbitrary. It is then a natural question whether the aforementioned Weyl-Feferman predicative mathematics suffers from the same compatibility problem. We show that this is the case in Section 4.2.

**1.3. Overview of main results.** Our main results, to be proved in detail, are that the *Cousin and Lindelöf lemmas* from Remark 1.1 live at the very edge of second-order arithmetic  $Z_2$ , in contrast to their countable substitutes (if existent) which live in very weak fragments of  $Z_2$ . The same result for the other theorems in Remark 1.1 readily follows from our main results, as discussed in Section 3.2.2.

Our results thus subvert the ‘Big Five’ picture of Reverse Mathematics from Section 2.1. We shall obtain our results by showing that it is *extremely* hard to compute the objects (like a finite sub-cover from the Cousin lemma) from the theorems in Remark 1.1. A precise statement is found at the end of this section.

**Remark 1.1** (Basic theorems).

- (i) *Cousin lemma*: **any** open cover of  $[0, 1]$  has a finite sub-cover ([13]).
- (ii) *Lindelöf lemma*: **any** open cover of  $\mathbb{R}$  has a countable sub-cover ([34]).
- (iii) *Besicovitch and Vitali covering lemmas* as in e.g. [1, §2].
- (iv) Basic properties of the gauge integral ([3]), like uniqueness, Hake’s theorem, and extension of the Riemann integral.
- (v) *Neighbourhood Function Principle NFP* ([57, p. 215]).
- (vi) The existence of *Lebesgue numbers* for **any** open cover ([24]).
- (vii) The *Banach-Alaoglu theorem* for **any** open cover ([55, X.2.4], [10, p. 140]).
- (viii) The *Heine-Young* and *Lusin-Young* theorems, the *tile theorem* [30, 62], and the latter’s generalisation due to Rademacher ([44, p. 190]).

According to Bourbaki’s historical note in [8, Ch. I], the by far most important ‘acquisition’ of Schoenflies’ monograph [52] is a theorem which constitutes a generalisation of the Cousin lemma. Another historical note is that Cousin (and Lindelöf in [34, p. 698]) talks about (uncountable) covers on [13, p. 22] in the following way:

*if to each point of  $S$  there corresponds a circle of finite radius ...*

In particular, any  $f : S \rightarrow \mathbb{R}^+$  gives rise to a cover *in the sense of the previous quote by Cousin* as follows:  $\cup_{x \in S} (x - f(x), x + f(x))$  covers  $S \subset \mathbb{R}$ .

We now make our main results precise, for which some definitions are needed. Detailed definitions may be found in Section 2.2.

**Definition 1.2.** Let  $Z_2$  be *second-order arithmetic* as defined in [55, I.2.4] and let  $\Pi_k^1\text{-CA}_0$  be the fragment of  $Z_2$  with comprehension restricted to  $\Pi_k^1$ -formulas.

As noted above, to formulate the theorems from the list, we require a richer language than that of  $Z_2$ . We shall make use of  $\text{RCA}_0^\omega$ , Kohlenbach's 'base theory' of higher-order Reverse Mathematics ([32, §2]), and the associated language of *all finite types*. We introduce this framework in full detail in Section 2.2.

**Definition 1.3.** Let  $(\exists^3)$  be the functional deciding set quantification from [7, p. 713], and let  $(S_k^2)$  similarly be the functional deciding  $\Pi_k^1$ -formulas; see Section 2.3 for the exact definition. For  $k = 1$ , the subscript is omitted and  $(S^2)$  is usually called the *Suslin functional*. Let  $\text{QF-AC}$  be the axiom of choice restricted to quantifier-free formulas as in [32, §1]. We define  $Z_2^\omega \equiv \text{RCA}_0^\omega + \text{QF-AC} + (\exists^3)$  and  $\Pi_k^1\text{-CA}_0^\omega \equiv \text{RCA}_0^\omega + \text{QF-AC} + (S_k^2)$ . The former (resp. the latter) system is the 'higher-type' analogue of  $Z_2$  (resp.  $\Pi_k^1\text{-CA}_0$ ).

Our main results from Section 3 are now as follows: (i)  $Z_2^\omega$  proves the Cousin and Lindelöf lemmas from the above list, (ii)  $\Pi_k^1\text{-CA}_0^\omega$  cannot prove any of the theorems in the above list, and (iii) the Cousin lemma is *equivalent* to basic properties of the gauge integral in Kohlenbach's aforementioned framework. In a nutshell, we show that *full second-order arithmetic*, in the form of  $(\exists^3)$ , is needed to prove the Cousin and Lindelöf lemmas, and the same for related theorems from Remark 1.1. Besides constituting a *huge* leap in logical strength (compared to  $\Pi_1^1\text{-CA}_0$ , the strongest Big Five system), the equivalences involving the gauge integral also seriously challenge the 'Big Five' classification from Reverse Mathematics (See Section 2.1).

Finally, as noted above, Reverse Mathematics is intimately connected to computability theory, and the same holds for our results; for instance, the functional defined by  $(\exists^3)$  (resp.  $(S_k^2)$ ) can (resp. cannot) compute (in the sense of Section 2.3) the finite sub-cover from the Cousin lemma on input an open cover of  $[0, 1]$ .

Our main results in Section 4 are then as follows: inspired by the aforementioned connection, we study the interaction between the theorems from the above list and the Big Five of Reverse Mathematics given by the Suslin functional  $S$  and Feferman's *search functional*  $\mu$  from Section 2.3. This leads to surprising results in (higher-order) Reverse Mathematics, as follows:

- (1) The combination of the Cousin lemma and Feferman's  $\mu$  yields *transfinite recursion* for arithmetical formulas, i.e. the fourth Big Five system. We derive novel theorems about Borel functions from this result.
- (2) The combination of the Cousin lemma and the Suslin functional  $S$  yields Gandy's *superjump*, the aforementioned 'higher-order' *Halting problem*.
- (3) The combination of the Lindelöf lemma *for Baire space* (given by a functional) and Feferman's  $\mu$  yields the Suslin functional  $S$ , i.e. the fifth Big Five system, and by the previous item also Gandy's superjump.

As will become clear in Section 4.2.2, the third item solves the compatibility problem of Weyl-Feferman predicativist mathematics from Section 1.2.3 in the negative. We also point out that the Lindelöf lemma (resp. the Cousin lemma) and Feferman's  $\mu$  are rather weak *in isolation*, and only become strong when combined.

## 2. PRELIMINARIES

We sketch the program *Reverse Mathematics* in Section 2.1 and discuss the associated framework *second-order arithmetic* in Section 2.2, as well as its generalisation *higher-order arithmetic*. As our main results will be proved using techniques from computability theory, we discuss some essential elements of the latter in Section 2.3.

**2.1. Introducing Reverse Mathematics.** Reverse Mathematics (RM) is a program in the foundations of mathematics initiated around 1975 by Friedman ([21, 22]) and developed extensively by Simpson ([55]) and others. We refer to [55] for an overview of RM and introduce the required definitions (like the ‘base theory’  $\text{RCA}_0$ ) in Section 2.2.1; we now sketch some of the aspects of RM essential to this paper.

The aim of RM is to find the axioms necessary to prove a statement of *ordinary*, i.e. *non-set theoretical* mathematics. The classical base theory  $\text{RCA}_0$  of ‘computable mathematics’, introduced in Section 2.2.1, is always assumed. Thus, the aim is:

*The aim of RM is to find the minimal axioms  $A$  such that  $\text{RCA}_0$  proves  $[A \rightarrow T]$  for statements  $T$  of ordinary mathematics.*

Surprisingly, once the minimal axioms  $A$  have been found, we almost always also have  $\text{RCA}_0 \vdash [A \leftrightarrow T]$ , i.e. not only can we derive the theorem  $T$  from the axioms  $A$  (the ‘usual’ way of doing mathematics), we can also derive the axiom  $A$  from the theorem  $T$  (the ‘reverse’ way of doing mathematics). In light of these ‘reversals’, the field was baptised ‘Reverse Mathematics’.

Perhaps even more surprisingly, in the majority of cases, for a statement  $T$  of ordinary mathematics, either  $T$  is provable in  $\text{RCA}_0$ , or the latter proves  $T \leftrightarrow A_i$ , where  $A_i$  is one of the logical systems  $\text{WKL}_0$ ,  $\text{ACA}_0$ ,  $\text{ATR}_0$  or  $\Pi_1^1\text{-CA}_0$ , which are introduced in the next section. The latter four together with  $\text{RCA}_0$  form the ‘Big Five’ and the aforementioned observation that most mathematical theorems fall into one of the Big Five categories, is called the *Big Five phenomenon* ([37, p. 432]). Furthermore, each of the Big Five has a natural formulation in terms of (Turing) computability (See [55, I]). As noted by Simpson in [55, I.12], each of the Big Five also corresponds (loosely) to a foundational program in mathematics.

Finally, we note that the Big Five systems of RM yield a linear order:

$$\Pi_1^1\text{-CA}_0 \rightarrow \text{ATR}_0 \rightarrow \text{ACA}_0 \rightarrow \text{WKL}_0 \rightarrow \text{RCA}_0. \quad (2.1)$$

By contrast, there are many incomparable *logical* statements in second-order arithmetic. For instance, a regular plethora of such statements may be found in the *Reverse Mathematics zoo* in [16]. The latter is intended as a collection of (somewhat natural) theorems outside of the Big Five classification of RM. However, the results sketched in Section 1.3 *fundamentally distort* the elegant picture (2.1).

**2.2. The framework of Reverse Mathematics.** We introduce axiomatic systems essential to RM. We start with a sketch of *second-order arithmetic* (See [55, I.2.4]), the framework of Friedman-Simpson RM, and finish with *higher-order arithmetic*, the framework of Kohlenbach’s *higher-order* RM (See [32]).

**2.2.1. Second-order arithmetic and fragments.** The language  $\mathcal{L}_2$  of second-order arithmetic  $\mathcal{Z}_2$  has two sorts of variables: *number variables*  $n, m, k, l, \dots$  intended to range over the natural numbers, and *set variables*  $X, Y, Z, \dots$  intended to range over sets of natural numbers. The constants of  $\mathcal{L}_2$  are  $0, 1, <_{\mathbb{N}}, +_{\mathbb{N}}, \times_{\mathbb{N}}, =_{\mathbb{N}}$  and  $\in$ ,

which are intended to have their usual meaning (by the axioms introduced below). Formulas and terms are built up from these constants in the usual way.

**Definition 2.1.** Second-order arithmetic  $Z_2$  consists of three axiom schemas:

- (1) Basic axioms expressing that  $0, 1, <_{\mathbb{N}}, +_{\mathbb{N}}, \times_{\mathbb{N}}$  form an ordered semi-ring with equality  $=_{\mathbb{N}}$ .
- (2) Induction: For any  $X$ ,  $(0 \in X \wedge (\forall n)(n \in X \rightarrow n+1 \in X)) \rightarrow (\forall n)(n \in X)$ .
- (3) Comprehension: For any formula  $\varphi(n)$  of  $L_2$  which does not involve the variable  $X$ , we have  $(\exists X)(\forall n)(n \in X \leftrightarrow \varphi(n))$ .

Induction is well-known, while comprehension intuitively expresses that any  $L_2$ -formula  $\varphi(n)$  yields a set  $X = \{n \in \mathbb{N} : \varphi(n)\}$  consisting of exactly those numbers  $n \in \mathbb{N}$  satisfying  $\varphi(n)$ . Now, *fragments* of  $Z_2$  are obtained by restricting comprehension (and induction), for which the following definition is needed.

**Definition 2.2.** [Formula classes]

- (1) A formula of  $L_2$  is *quantifier-free* ( $\Sigma_0^0$  or  $\Pi_0^0$ ) if it does not involve quantifiers. To be clear: variables are allowed; only quantifiers are banned.
- (2) A formula of  $L_2$  is *arithmetical* ( $\Sigma_0^1$  or  $\Pi_0^1$ ) if it only involves quantifiers over number variables, i.e. set quantifiers like  $(\exists X)$  and  $(\forall Y)$  are not allowed.
- (3) An arithmetical formula is  $\Sigma_{k+1}^0$  (resp.  $\Pi_{k+1}^0$ ) if it has the form  $(\exists n)\varphi(n)$  (resp.  $(\forall n)\varphi(n)$ ) with  $\varphi$  in  $\Pi_k^0$  (resp. in  $\Sigma_k^0$ ).
- (4) A formula of  $L_2$  is  $\Sigma_{k+1}^1$  (resp.  $\Pi_{k+1}^1$ ) if it has the form  $(\exists X)\varphi(X)$  (resp.  $(\forall X)\varphi(X)$ ) with  $\varphi$  in  $\Pi_k^1$  (resp. in  $\Sigma_k^1$ ).
- (5) A formula of  $L_2$  is  $\Delta_{k+1}^i$  if it is both  $\Pi_{k+1}^i$  and  $\Sigma_{k+1}^i$  for  $i = 0, 1$ .

Intuitively, a  $\Sigma_k^0$ -formula is a quantifier-free formula pre-fixed by  $k$  alternating number quantifiers, starting with an existential one; a  $\Sigma_k^1$ -formula is an arithmetical formula pre-fixed by  $k$  alternating set quantifiers, starting with an existential one. The  $\Pi$ -formulas are (equivalent to) negations of the corresponding  $\Sigma$ -versions.

Using the above, the third and fifth ‘Big Five’ systems  $ACA_0$  and  $\Pi_1^1\text{-CA}_0$  are just  $Z_2$  with comprehension restricted to resp. arithmetical and  $\Pi_1^1$ -formulas. Alternatively,  $ACA_0$  allows one to build sets using *finite* iterations of Turing’s *Halting problem* ([58]), aka the *Turing jump*; intuitively,  $ATR_0$  extends this to *transfinite recursion*, i.e. the *unbounded* iteration of the Turing jump *along any countable well-ordering*. Furthermore, the ‘base theory’  $RCA_0$  is  $Z_2$  with comprehension restricted to  $\Delta_1^0$ -formulas, plus induction for  $\Sigma_1^0$ -formulas. As discussed in [55, II and IX.3],  $\Delta_1^0$ -comprehension essentially expresses that ‘all computable sets exists’, while  $\Sigma_1^0$ -induction corresponds to primitive recursion in the sense of Hilbert’s *finitistic mathematics*. The system  $WKL_0$  is just  $RCA_0$  extended by the *weak König’s lemma* which states that an infinite binary tree has a path.

Finally, in light of the previous and (2.1), the Big Five only constitute a *very tiny fragment* of  $Z_2$ ; on a related note, the RM of topology does give rise to theorems equivalent to  $\Pi_2^1\text{-CA}_0$  ([39]), but that is the current upper bound of RM to the best of our knowledge. In particular, if  $\Pi_k^1\text{-CA}_0$  is  $Z_2$  restricted to  $\Pi_k^1$ -comprehension, then this system can be said to ‘go beyond Friedman-Simpson RM’ for  $k \geq 3$ .

**2.2.2. Higher-order arithmetic and fragments.** As suggested by its name, *higher-order arithmetic* extends second-order arithmetic. Indeed, while the latter is restricted to numbers and sets of numbers, higher-order arithmetic also has sets of

sets of numbers, sets of sets of sets of numbers, et cetera. To formalise this idea, we introduce the collection of *all finite types*  $\mathbf{T}$ , defined by the two clauses:

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

where  $0$  is the type of natural numbers, and  $\sigma \rightarrow \tau$  is the type of mappings from objects of type  $\sigma$  to objects of type  $\tau$ . In this way,  $1 \equiv 0 \rightarrow 0$  is the type of functions from numbers to numbers, and where  $n + 1 \equiv n \rightarrow 0$ . Viewing sets as given by their characteristic function, we note that  $\mathbf{Z}_2$  only includes objects of type  $0$  and  $1$ .

The language of  $\mathbf{L}_\omega$  consists of variables  $x^\rho, y^\rho, z^\rho, \dots$  of any finite type  $\rho \in \mathbf{T}$ . Types may be omitted when they can be inferred from context. The constants of  $\mathbf{L}_\omega$  includes the type  $0$  objects  $0, 1$  and  $<_0, +_0, \times_0, =_0$  which are intended to have the same meaning as their  $\mathbb{N}$ -subscript counterparts in  $\mathbf{Z}_2$ . Equality at higher types is defined in terms of ‘ $=_0$ ’ as follows: for any objects  $x^\tau, y^\tau$ , we have

$$[x =_\tau y] \equiv (\forall z_1^{\tau_1} \dots z_k^{\tau_k}) [xz_1 \dots z_k =_0 yz_1 \dots z_k], \quad (2.2)$$

if the type  $\tau$  is composed as  $\tau \equiv (\tau_1 \rightarrow \dots \rightarrow \tau_k \rightarrow 0)$ . Furthermore,  $\mathbf{L}_\omega$  also includes the *recursor constant*  $\mathbf{R}_\sigma$  for any  $\sigma \in \mathbf{T}$ , which allows for iteration on type  $\sigma$ -objects as in the special case (2.3). Formulas and terms are defined as usual.

**Definition 2.3.** The base theory  $\mathbf{RCA}_0^\omega$  consists of the following axioms:

- (1) Basic axioms expressing that  $0, 1, <_0, +_0, \times_0$  form an ordered semi-ring with equality  $=_0$ .
- (2) Basic axioms defining the well-known  $\Pi$  and  $\Sigma$  combinators (aka  $K$  and  $S$  in [2]), which allow for the definition of  $\lambda$ -abstraction.
- (3) The defining axiom of the recursor constant  $\mathbf{R}_0$ : For  $m^0$  and  $f^1$ :

$$\mathbf{R}_0(f, m, 0) := m \text{ and } \mathbf{R}_0(f, m, n + 1) := f(\mathbf{R}_0(f, m, n)). \quad (2.3)$$

- (4) The *axiom of extensionality*: for all  $\rho, \tau \in \mathbf{T}$ , we have:

$$(\forall x^\rho, y^\rho, \varphi^{\rho \rightarrow \tau}) [x =_\rho y \rightarrow \varphi(x) =_\tau \varphi(y)]. \quad (\mathbf{E}_{\rho, \tau})$$

- (5) The induction axiom for quantifier-free<sup>4</sup> formulas.
- (6)  $\mathbf{QF-AC}^{1,0}$ : The quantifier-free axiom of choice as in Definition 2.4.

**Definition 2.4.** The axiom  $\mathbf{QF-AC}$  consists of the following for all  $\sigma, \tau \in \mathbf{T}$ :

$$(\forall x^\sigma)(\exists y^\tau)A(x, y) \rightarrow (\exists Y^{\sigma \rightarrow \tau})A(x, Y(x)) \quad (\mathbf{QF-AC}^{\sigma, \tau})$$

for any quantifier-free formula in the language of  $\mathbf{L}_\omega$ .

As discussed in [32, §2],  $\mathbf{RCA}_0^\omega$  and  $\mathbf{RCA}_0$  prove the same sentences ‘up to language’ as the latter is set-based and the former function-based. Recursion as in (2.3) is called *primitive recursion*; the class of functionals obtained from  $\mathbf{R}_\rho$  for all  $\rho \in \mathbf{T}$  is called *Gödel’s system T* of all (higher-order) primitive recursive functionals.

We use the usual notations for natural, rational, and real numbers, and the associated functions, as introduced in [32, p. 288-289].

**Definition 2.5** (Real numbers and related notions in  $\mathbf{RCA}_0^\omega$ ).

- (1) Natural numbers correspond to type zero objects, and we use ‘ $n^0$ ’ and ‘ $n \in \mathbb{N}$ ’ interchangeably. Rational numbers are defined as signed quotients of natural numbers, and ‘ $q \in \mathbb{Q}$ ’ and ‘ $<_{\mathbb{Q}}$ ’ have their usual meaning.

<sup>4</sup>To be absolutely clear, similar to Definition 2.2, variables (of any finite type) are allowed in quantifier-free formulas: only quantifiers are banned.



- (2) Real numbers are coded by fast-converging Cauchy sequences  $q_{(\cdot)} : \mathbb{N} \rightarrow \mathbb{Q}$ , i.e. such that  $(\forall n^0, i^0)(|q_n - q_{n+i}| <_{\mathbb{Q}} \frac{1}{2^n})$ . We use Kohlenbach's 'hat function' from [32, p. 289] to guarantee that every  $f^1$  defines a real number.
- (3) We write ' $x \in \mathbb{R}$ ' to express that  $x^1 := (q_{(\cdot)}^1)$  represents a real as in the previous item and write  $[x](k) := q_k$  for the  $k$ -th approximation of  $x$ .
- (4) Two reals  $x, y$  represented by  $q_{(\cdot)}$  and  $r_{(\cdot)}$  are *equal*, denoted  $x =_{\mathbb{R}} y$ , if  $(\forall n^0)(|q_n - r_n| \leq \frac{1}{2^{n-1}})$ . Inequality ' $<_{\mathbb{R}}$ ' is defined similarly.
- (5) Functions  $F : \mathbb{R} \rightarrow \mathbb{R}$  mapping reals to reals are represented by  $\Phi^{1 \rightarrow 1}$  mapping equal reals to equal reals, i.e.  $(\forall x, y)(x =_{\mathbb{R}} y \rightarrow \Phi(x) =_{\mathbb{R}} \Phi(y))$ .
- (6) The relation ' $x \leq_{\tau} y$ ' is defined as in (2.2) but with ' $\leq_0$ ' instead of ' $=_0$ '.
- (7) Sets of type  $\rho$  objects  $X^{\rho \rightarrow 0}, Y^{\rho \rightarrow 0}, \dots$  are given by their characteristic functions  $f_X^{\rho \rightarrow 0}$ , i.e.  $(\forall x^{\rho})[x \in X \leftrightarrow f_X(x) =_0 1]$ , where  $f_X^{\rho \rightarrow 0} \leq_{\rho \rightarrow 0} 1$ .

We sometimes omit the subscript ' $\mathbb{R}$ ' if it is clear from context. Finally, we introduce some notation to handle finite sequences nicely.

**Notation 2.6** (Finite sequences). We assume a dedicated type for 'finite sequences of objects of type  $\rho$ ', namely  $\rho^*$ . Since the usual coding of pairs of numbers goes through in  $\text{RCA}_0^{\omega}$ , we shall not always distinguish between 0 and  $0^*$ . Similarly, we do not always distinguish between ' $s^{\rho}$ ' and ' $\langle s^{\rho} \rangle$ ', where the former is 'the object  $s$  of type  $\rho$ ', and the latter is 'the sequence of type  $\rho^*$  with only element  $s^{\rho}$ '. The empty sequence for the type  $\rho^*$  is denoted by ' $\langle \rangle_{\rho}$ ', usually with the typing omitted.

Furthermore, we denote by ' $|s| = n$ ' the length of the finite sequence  $s^{\rho^*} = \langle s_0^{\rho}, s_1^{\rho}, \dots, s_{n-1}^{\rho} \rangle$ , where  $|\langle \rangle| = 0$ , i.e. the empty sequence has length zero. For sequences  $s^{\rho^*}, t^{\rho^*}$ , we denote by ' $s*t$ ' the concatenation of  $s$  and  $t$ , i.e.  $(s*t)(i) = s(i)$  for  $i < |s|$  and  $(s*t)(j) = t(|s| - j)$  for  $|s| \leq j < |s| + |t|$ . For a sequence  $s^{\rho^*}$ , we define  $\bar{s}N := \langle s(0), s(1), \dots, s(N-1) \rangle$  for  $N^0 < |s|$ . For a sequence  $\alpha^{0 \rightarrow \rho}$ , we also write  $\bar{\alpha}N = \langle \alpha(0), \alpha(1), \dots, \alpha(N-1) \rangle$  for *any*  $N^0$ . By way of shorthand,  $(\forall q^{\rho} \in Q^{\rho^*})A(q)$  abbreviates  $(\forall i^0 < |Q|)A(Q(i))$ , which is (equivalent to) quantifier-free if  $A$  is.

**2.3. Higher-order computability theory.** As noted above, our main results will be proved using techniques from computability theory. Thus, we first make our notion of 'computability' precise as follows.

- (I) We adopt ZFC, i.e. Zermelo-Fraenkel set theory with the Axiom of Choice, 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 [35, 49]) as our official notion of 'computable'.

For the rest of this section, we introduce some existing functionals which will be used below. These functionals constitute the counterparts of  $Z_2$ , and some of the Big Five systems, in higher-order RM. First of all,  $\text{ACA}_0$  is readily derived from:

$$(\exists \mu^2)[(\forall f^1)((\exists n)(f(n) = 0 \rightarrow f(\mu(f)) = 0)], \quad (\mu^2)$$

and  $\text{ACA}_0^{\omega} \equiv \text{RCA}_0^{\omega} + \text{QF-AC} + (\mu^2)$  proves the same  $\Pi_2^1$ -sentences as  $\text{ACA}_0$  by [48, Theorem 2.2]. The (unique) functional  $\mu^2$  in  $(\mu^2)$  is also called *Feferman's  $\mu$*  ([2]), and is clearly *discontinuous* at  $f =_1 11\dots$ ; in fact,  $(\mu^2)$  is equivalent to the existence of  $F : \mathbb{R} \rightarrow \mathbb{R}$  such that  $F(x) = 1$  if  $x >_{\mathbb{R}} 0$ , and 0 otherwise ([32, §3]).

Secondly,  $\Pi_1^1\text{-CA}_0$  is readily derived from the following sentence:

$$(\exists S^2)(\forall f^1)[(\exists g^1)(\forall x^0)(f(\bar{g}n) = 0) \leftrightarrow S(f) = 0], \quad (S^2)$$

and  $\Pi_1^1\text{-CA}_0^\omega \equiv \text{RCA}_0^\omega + \text{QF-AC} + (S^2)$  proves the same  $\Pi_3^1$ -sentences as  $\Pi_1^1\text{-CA}_0$  by [48, Theorem 2.2]. The (unique) functional  $S^2$  in  $(S^2)$  is also called *the Suslin functional* ([32]). By definition, the Suslin functional  $S^2$  can decide whether a  $\Sigma_1^1$ -formula (as in the left-hand side of  $(S^2)$ ) is true or false. We similarly define the functional  $S_k^2$  which decides the truth or falsity of  $\Sigma_k^1$ -formulas; we also define the system  $\Pi_k^1\text{-CA}_0^\omega$  as  $\text{RCA}_0^\omega + \text{QF-AC} + (S_k^2)$ , where  $(S_k^2)$  expresses that  $S_k^2$  exists.

Thirdly, full second-order arithmetic  $Z_2$  is readily derived from the sentence:

$$(\exists E^3)(\forall Y^2)[(\exists f^1)Y(f) = 0 \leftrightarrow E(Y) = 0], \quad (\exists^3)$$

and we define  $Z_2^\omega \equiv \text{RCA}_0^\omega + \text{QF-AC} + (\exists^3)$ . The (unique) functional from  $(\exists^3)$  is also called ‘ $\exists^3$ ’, and we will use a similar convention for other functionals.

Fourth, there is primitive recursive function  $U$  such that ‘ $U(e, k, n) =_0 m + 1$ ’ expresses that the  $e$ -th Turing machine with input  $k$  halts after  $n$  steps with output  $m$ . By definition, Feferman’s  $\mu^2$  provides an upper bound on this  $n$  if it exists, i.e. we can use  $\mu^2$  to solve the Halting problem. Similarly, Gandy’s *superjump* solves the Halting problem for higher-order computability as follows:

$$\mathbb{S}(F^2, e^0) := \begin{cases} 0 & \text{if } \{e\}(F) \text{ terminates} \\ 1 & \text{otherwise} \end{cases}, \quad (\mathbb{S}^3)$$

where  $e$  is an S1-S9-index. A characterisation of  $\mathbb{S}$  in terms of discontinuities may be found in [26]. Hence, there are a number of similarities between  $\mu^2$  and  $\mathbb{S}$ .

Finally, recall that the Cousin lemma from Remark 1.1 states the existence of a finite sub-cover for an open cover of the unit interval. Since Cantor space is homeomorphic to a closed subset of  $[0, 1]$ , the former inherits the same property. In particular, for any  $G^2$ , the corresponding ‘canonical cover’ of  $2^\mathbb{N}$  is  $\cup_{f \in 2^\mathbb{N}} [\bar{f}G(f)]$  where  $[\sigma^{0^*}]$  is the set of all binary extensions of  $\sigma$ . By compactness, there is a finite sequence  $\langle f_0, \dots, f_n \rangle$  such that the set of  $\cup_{i \leq n} [\bar{f}_i F(f_i)]$  still covers  $2^\mathbb{N}$ . We now introduce the specification  $\text{SCF}(\Theta)$  for a (non-unique) functional  $\Theta$  which computes such a finite sequence. We refer to such a functional  $\Theta$  as a *realiser* for the compactness of Cantor space, and simplify its type to ‘3’ to improve readability.

**Definition 2.7.** The formula  $\text{SCF}(\Theta)$  is as follows for  $\Theta^{2 \rightarrow 1^*}$ :

$$(\forall G^2)(\forall f^1 \leq_1 1)(\exists g \in \Theta(G))(f \in [\bar{g}G(g)]). \quad (2.4)$$

where ‘ $f \in [\bar{g}G(g)]$ ’ is the quantifier-free formula  $\bar{f}G(g) =_{0^*} \bar{g}G(g)$ .

Clearly, there is no unique  $\Theta$  as in (2.4) (just add more binary sequences to  $\Theta(G)$ ); nonetheless, we have in the past referred to any  $\Theta$  satisfying  $\text{SCF}(\Theta)$  as ‘the’ *special fan functional*  $\Theta$ , and we will continue this abuse of language. We shall however repeatedly point out the non-unique nature of the special fan functional  $\Theta$  in the following. While  $\Theta$  may appear exotic at first, it provides the only method we can think of for computing gauge integrals *in general*, as discussed in Remark 3.24.

Finally, as to its provenance,  $\Theta$  was introduced as part of the study of the *Gandy-Hyland functional* in [50, §2] via a slightly different definition. These definitions are identical up to a term of Gödel’s  $T$  of low complexity.

## 3. MAIN RESULTS I

We establish our main results as sketched in Section 1.3. We treat the Cousin lemma in full detail in Section 3.1, while similar ‘covering theorems’ from Remark 1.1 are treated analogously in Section 3.2. We show in Section 3.3 that the Cousin lemma is equivalent to various basic properties of the *gauge integral*.

**3.1. Cousin lemma.** Cousin first proved (what is now known as) the *Cousin lemma* before 1893 ([15]). This lemma implies that  $I = [0, 1]$  is Heine-Borel compact, i.e. that any open cover of  $I$  has a finite sub-cover. The goal of this section is to establish that, despite its seemingly elementary nature, the Cousin lemma can only be proved in full second-order arithmetic, as sketched in Section 1.3. This should be contrasted with the restriction to *countable* covers, which may be proved in the weak fragment  $\text{WKL}_0$  of second-order arithmetic by [55, IV.1.2]).

First of all, a functional  $\Psi : \mathbb{R} \rightarrow \mathbb{R}^+$  gives rise to the (uncountable) *canonical* open cover  $\cup_{x \in I} I_x^\Psi$  where  $I_x^\Psi$  is the open interval  $(x - \Psi(x), x + \Psi(x))$ . Hence, the Cousin lemma implies that  $\cup_{x \in I} I_x^\Psi$  has a finite sub-cover; in symbols:

$$(\forall \Psi : \mathbb{R} \rightarrow \mathbb{R}^+)(\exists \langle y_1, \dots, y_k \rangle)(\forall x \in I)(\exists i \leq k)(x \in I_{y_i}^\Psi). \quad (\text{HBU})$$

The main goal of this section is now to prove the following theorem, which establishes that full second-order arithmetic is needed to prove the Cousin lemma.

**Theorem 3.1.** *The system  $Z_2^\omega$  proves HBU; no system  $\Pi_k^1\text{-CA}_0^\omega$  proves HBU.*

The first part is a necessity as otherwise the aforementioned designation ‘analysis’ for  $Z_2$  would be meaningless. The second part constitutes a surprise: the restriction of HBU to *countable* covers is equivalent to  $\text{WKL}_0$  ([55, IV.1]), a system with the (first-order) strength of  $\text{RCA}_0^\omega$ . Kohlenbach has introduced generalisations of  $\text{WKL}_0$  with properties similar to HBU ([33, §5-6]), but these axioms *do not stem from mathematics*, i.e. are ‘purely logical’. Furthermore, HBU is *robust* ([37, p. 432]) in that restricting the variable  $x$  to the (Turing) computable reals or the rationals in  $I$  does not make a difference. We now prove the first part of Theorem 3.1.

**Theorem 3.2.** *The system  $Z_2^\omega$  proves HBU.*

*Proof.* We only sketch the proof as it makes use of items from Remark 1.1 to be studied in Section 3.2. A full proof may be found in Theorem 3.14. Now, to derive HBU, we note that the *Lindelöf lemma* provides a *countable* sub-cover for any open cover of  $I$ . Since  $(\exists^3)$  immediately implies  $Z_2$ , we may use [55, IV.1.2], which implies that every countable open cover has a finite sub-cover. What remains is to prove the Lindelöf lemma, which readily follows from the *Neighbourhood function principle* NFP, i.e. item (v) in Remark 1.1, as will become clear in the proof of Theorem 3.14. In turn, NFP has a straightforward proof in  $Z_2^\omega$ , as will also become clear in the proof of Theorem 3.14.  $\square$

As noted above, we shall make use of computability theory to establish Theorem 3.1. Hence, we first show that HBU is equivalent to the existence of the special fan functional  $\Theta$  in Theorem 3.3. Theorem 3.1 will then be established by showing that models of  $\Pi_k^1\text{-CA}_0^\omega$  do not always contain  $\Theta$  as in Theorem 3.4. Note that the functional  $\Omega$  as in (3.1) is called a *realiser* for HBU.

**Theorem 3.3.**  *$\text{ACA}_0^\omega$  proves that  $(\exists \Theta)\text{SCF}(\Theta)$  is equivalent to HBU and to*

$$(\exists \Omega^{2 \rightarrow 1^*})(\forall \Psi : \mathbb{R} \rightarrow \mathbb{R}^+)(\forall x \in [0, 1])(\exists y \in \Omega(\Psi)(x \in I_y^\Psi)). \quad (3.1)$$

*Proof.* We first point out two useful properties of Feferman's  $\mu$ : the axiom ( $\mu^2$ ) defining the latter functional is equivalent to the existence of  $F : \mathbb{R} \rightarrow \mathbb{R}$  such that  $F(x) = 1$  if  $x >_{\mathbb{R}} 0$ , and 0 otherwise ([32, §3]). Furthermore, by repeatedly applying  $\mu$ , we can show that any arithmetical formula is equivalent to a quantifier-free one. We also recall the notation ' $f \in [\sigma]$ ' for covers of Cantor space from Definition 2.7.

Based on the previous, given  $\Psi, y_1, \dots, y_k$  as in HBU, we can decide if the intervals  $I_{y_i}^\Psi$  form an open covering or not: we just check (using  $\mu$ ) how the end-points of these intervals are interleaved. Thus, using  $\mu$  as a parameter, we can deduce (3.1) from HBU by QF-AC. Likewise, given  $f_1, \dots, f_n \leq 1$  and  $k_1, \dots, k_n$  in  $\mathbb{N}$ , we can decide if the set of neighbourhoods  $[\bar{f}_i, k_i]$  form a covering or not; hence, we may use QF-AC to obtain  $\Theta$  from the compactness of Cantor space in the same way.

Now define  $\xi(f) = \sum_{i \in \mathbb{N}} f(i) \cdot 2^{-(i+1)}$  and  $\zeta(f) = \sum_{i \in \mathbb{N}} 2f(i) \cdot 3^{-(i+1)}$  for  $f \in \{0, 1\}^{\mathbb{N}}$ ; note that  $\xi$  is a continuous projection of  $\{0, 1\}^{\mathbb{N}}$  to  $[0, 1]$ , while  $\zeta$  is the homeomorphism between  $\{0, 1\}^{\mathbb{N}}$  and the classical Cantor space  $C^c$ . Using  $\xi$  and  $\zeta$ , we can convert canonical covers between  $I$  and Cantor space as follows:

- For  $\Psi : [0, 1] \rightarrow \mathbb{R}^+$ , define  $F_\Psi(f)$  as the least  $n$  such that  $[\bar{f}n] \subseteq \xi^{-1}(I_{\xi(f)}^\Psi)$ .
- For  $F : \{0, 1\}^{\mathbb{N}} \rightarrow \mathbb{N}$ , we define  $\Psi_F(x)$  as the distance from  $x$  to  $C^c$  if  $x \notin C^c$ , and as the least rational (in some canonical enumeration of  $\mathbb{Q}^+$ )  $q$  such that  $\zeta^{-1}((x - q, x + q)) \subseteq [\zeta^{-1}(x)F(\zeta^{-1}(x))]$  if  $x \in C^c$ .

These constructions are arithmetical, and the compactness property for the associated coverings are transferred from one space to the other in both directions.  $\square$

From the proof, we may also conclude that there is a term  $t$  such that if  $\text{SCF}(\Theta)$  and  $\Omega := t(\Theta, \mu)$  then  $\Omega$  satisfies (3.1), and conversely, there is a term  $s$  such that if  $\Omega$  satisfies (3.1) and  $\Theta := s(\Omega, \mu)$ , then  $\text{SCF}(\Theta)$ . The proof makes use of the *Axiom of Choice* (as in QF-AC) to obtain a functional  $\Theta$  as in  $\text{SCF}(\Theta)$ , resp.  $\Omega$  satisfying (3.1), from the existence of finite sub-coverings. Nonetheless, a careful analysis of the known proofs of HBU yields such functionals  $\Theta$  and  $\Omega$  *without the Axiom of Choice*. We discuss this in more detail in Remark 3.9 at the end of this section.

To establish Theorem 3.1, we now exhibit a model (aka type structure) of  $\Pi_k^1\text{-CA}_0^\omega$  in which there is no special fan functional; HBU is then false in that model by Theorem 3.3, and hence  $\Pi_k^1\text{-CA}_0^\omega$  cannot prove HBU by the soundness theorem.

**Theorem 3.4.** *There is a type structure validating  $\Pi_k^1\text{-CA}_0^\omega$  for all  $k$  and at the same time satisfying  $(\forall\Theta^3)\text{-SCF}(\Theta)$ .*

*Proof.* We introduce a family of type structures validating  $(\forall\Theta^3)\text{-SCF}(\Theta)$ . Theorem 3.8 below tells us that one of those structures contains all  $S_k^1$  and is closed under S1-S9, establishing the theorem. Intuitively speaking, we start from a  $\beta$ -model  $A$  and have that any functional  $G : A \rightarrow \mathbb{N}$  which is computable in some  $S_k^2$  and elements from  $A$  will be total over  $\mathbb{N}^{\mathbb{N}}$  by the same algorithm. By absoluteness, there are  $f_1, \dots, f_n$  in  $A$  inducing a covering of  $2^{\mathbb{N}}$  of the standard form. Since it is flexible which objects of type 2 we include in an extension of  $A$  to a typed structure,  $A$  together with the  $S_k^2$ 's cannot "decide" whether there is  $\Theta$  as in  $\text{SCF}(\Theta)$ .

Let  $A \subseteq \mathbb{N}^{\mathbb{N}}$  be a countable set such that all  $\Pi_k^1$ -statements with parameters from  $A$  are absolute for  $A$ . Also, let  $S_k^2$  be the characteristic function of a complete  $\Pi_k^1$ -set for each  $k$ ; we also write  $S_k^2$  for the restriction of this functional to  $A$ . Clearly, for  $f \in \mathbb{N}^{\mathbb{N}}$  computable in any  $S_k^2$  and some  $f_1, \dots, f_n$  from  $A$ ,  $f$  is also in  $A$ .

**Convention 3.5.** Since  $A$  is countable, we write  $A$  as the increasing union  $\bigcup_{k \in \mathbb{N}} A_k$  where  $A_0$  consists of the hyperarithmetical functions and for  $k > 0$  we have:

- There is an element in  $A_k$  enumerating  $A_{k-1}$ .
- $A_k$  is the closure of a finite set  $g_1, \dots, g_{n_k}$  under computability in  $S_k^2$ .

For the sake of uniform terminology, we rename  $\exists^2$  to  $S_0^2$  and let the associated finite sequence  $g_1, \dots, g_{n_0}$  be the empty list.

We now define the functional  $F^2$  on  $A$  as follows.

**Definition 3.6.** [The functional  $F$ ] Define  $F(f)$  for  $f \in A$  as follows:

- If  $f \notin 2^{\mathbb{N}}$ , put  $F(f) := 0$ .
- If  $f \in 2^{\mathbb{N}}$ , let  $k$  be minimal such that  $f \in A_k$ . We put  $F(f) := 2^{-(k+2+e)}$ , where  $e$  is a ‘minimal’ index for computing  $f$  from  $S_k^2$  and  $\{g_1, \dots, g_{n_k}\}$  as follows: the ordinal rank of this computation of  $f$  is minimal and  $e$  is minimal among the indices for  $f$  of the same ordinal rank.

By definition,  $F$  as in Definition 3.6 is injective on  $A_0$  and on each set  $A_{k+1} \setminus A_k$ . Moreover, if  $\mathbf{m}$  is the usual measure on  $2^{\mathbb{N}}$ , we see that

$$\mathbf{m}\left(\bigcup_{f \in A_0} [\bar{f}F(f)]\right) \leq 2^{-1} \text{ and } \mathbf{m}\left(\bigcup_{f \in A_{k+1} \setminus A_k} [\bar{f}F(f)]\right) \leq 2^{-(k+2)}.$$

As a consequence, if  $F$  is extended to a total functional  $G$  and  $\Theta$  satisfies SCF( $\Theta$ ), then  $\Theta(G)$  cannot be a finite list from  $A$ . So, for any type structure  $\mathbf{Tp} = \{\mathbf{Tp}_n\}_{n \in \mathbb{N}}$  where  $\mathbf{Tp}_0 = \mathbb{N}$ ,  $\mathbf{Tp}_1 = A$  and  $F \in \mathbf{Tp}_2$ , there is no instance of  $\Theta$  as in SCF( $\Theta$ ) in  $\mathbf{Tp}_3$ . To establish the theorem, we require one such type structure, containing each  $S_k^2$  and  $F$ , and closed under Kleene’s S1 - S9; such a type structure is provided by Theorem 3.8, i.e. the latter establishes the theorem, and we are done.  $\square$

For Theorem 3.8, we require some properties of  $F^2$  from Definition 3.6.

**Lemma 3.7** (Properties of the functional  $F$ ).

- (1) For each  $k$ , the restriction of  $F$  to  $A_k$  is computable in the functions  $g_1, \dots, g_{n_k}$  from Convention 3.5, and the functional  $S_k^2$ .
- (2) Let  $G$  be any total extension of  $F$ , let  $f_1, \dots, f_m \in A$ , and assume that the function  $f$  is computable in  $G$ ,  $f_1, \dots, f_m$  and some  $S_k^2$ . Then also  $f \in A$ .

*Proof.* For the first part, we use induction on  $k$ . For  $k = 0$ , we use *Gandy selection* ([35, p. 210]) for  $\exists^2$  which permits us to compute an  $\exists^2$  index for each hyperarithmetical function. For  $k > 0$ , we use that  $S_l^2$  is computable in  $S_k^2$  when  $l < k$  and that we have enumerations of each of the sets  $A_0, \dots, A_{k-1}$  computable in  $g_1, \dots, g_{n_k}$  and  $S_k^2$ . Then we can apply the induction hypothesis for  $f \in A_l$  for some  $l < k$  and the Gandy selection method relative to  $S_k^2$  for  $f \in A_k \setminus A_{k-1}$ . For the second part, without loss of generality, we may assume that  $f_1, \dots, f_m$  are all in  $A_k$ . By the first part of this lemma,  $G$  restricted to  $A_k$  is computable in  $S_k^2$ , and  $A_k$  is closed under computations relative to  $S_k^2$ . The claim now follows.  $\square$

**Theorem 3.8.** *There is a type structure  $\{\mathbf{Tp}_n\}_{n \in \mathbb{N}}$ , closed under Kleene’s S1-S9, such that  $\mathbf{Tp}_0 = \mathbb{N}$  and:*

- (1)  $\mathbf{Tp}_1$  is a countable subset  $A$  of  $\mathbb{N}^{\mathbb{N}}$  such that all analytical statements (i.e. any  $\Pi_m^1$ -sentence, for any  $m$ ) are absolute for  $A$ .
- (2)  $\mathbf{Tp}_2$  contains the restrictions of all  $S_k^2$  to  $A$ .
- (3) There exists  $F \in \mathbf{Tp}_2$  inducing an open covering of  $A$  for which there is no finite sub-covering in the type structure.

*Proof.* The theorem expresses exactly what  $\text{Tp}_n$  has to be for  $n = 0$  and  $n = 1$ . For  $n > 1$  we recursively let  $\text{Tp}_n$  consist of all functionals  $\phi : \text{Tp}_{n-1} \rightarrow \mathbb{N}$  that are S1-S9-computable in  $F$ , some  $S_k^2$ , and elements from  $A$ , where  $F$  is as in Definition 3.6. This type structure has the desired property.  $\square$

The proof of Theorem 3.1 is now done. We finish this section with a remark.

**Remark 3.9** (The Axiom of Choice and  $\Theta$ ). First of all, while the (quantifier-free) Axiom of Choice is used to establish the existence of  $\Theta$  in Theorem 3.3, the special fan functional  $\Theta$  can be computed from  $\exists^3$  (via a term from Gödel's  $T$ ) *without any* use of the Axiom of Choice by [51, Cor. 3.29]. Furthermore, Borel's construction from [9, p. 52] can be applied to our notion of canonical cover, yielding a *countable sub-cover* without using the Axiom of Choice. Furthermore, the instance  $\Theta_0$  of the special fan functional from [42, §5.1] is defined using Borel's construction.

**3.2. Lindelöf lemma and similar theorems.** We establish results analogous to Theorem 3.1 for some of the other theorems from Remark 1.1. We discuss how these theorems are used in mathematics in Remark 3.16.

**3.2.1. Lindelöf lemma.** We recall that Lindelöf proved the *Lindelöf lemma* in 1903 ([34]) which states that any open cover of any subset of  $\mathbb{R}^n$  has a countable sub-cover. In this section, we study this lemma restricted to  $\mathbb{R}$ , while Baire space is studied in Section 4.2.2. We are interested in the following formulations.

**Definition 3.10.** [LIND] For every  $\Psi : \mathbb{R} \rightarrow \mathbb{R}^+$ , there is a sequence of open intervals  $\cup_{n \in \mathbb{N}} (a_n, b_n)$  covering  $\mathbb{R}$  such that  $(\forall n \in \mathbb{N})(\exists x \in \mathbb{R})[(a_n, b_n) = I_x^\Psi]$ .

**Definition 3.11.** [LIND<sub>2</sub>]  $(\forall \Psi : \mathbb{R} \rightarrow \mathbb{R}^+)(\exists \Phi^{0 \rightarrow 1})(\forall x \in \mathbb{R})(\exists n^0)(x \in I_{\Phi(n)}^\Psi)$ .

**Definition 3.12.** [LIND<sub>3</sub>]  $(\exists \Xi)(\forall \Psi : \mathbb{R} \rightarrow \mathbb{R}^+)(\forall x \in \mathbb{R})(\exists n^0)(x \in I_{\Xi(\Psi)(n)}^\Psi)$ .

The following theorem establishes the connection between LIND and HBU, while also showing that the introduction of  $\Xi$  does not change LIND much.

**Theorem 3.13.** *The system  $\text{RCA}_0^\omega$  proves  $[\text{LIND} + \text{WKL}_0] \leftrightarrow \text{HBU}$  and  $\text{ACA}_0^\omega$  proves  $\text{LIND} \leftrightarrow \text{LIND}_2 \leftrightarrow \text{LIND}_3$ .*

*Proof.* For the first part,  $\text{WKL}_0$  implies that every *countable* cover of  $I$  has a finite sub-cover by [55, IV.1.2]. Hence,  $\text{LIND} + \text{WKL}_0 \rightarrow \text{HBU}$  is immediate, while HBU clearly generalises to  $[-N, N]$  for any natural number  $N^0$ . Putting all the finite sub-covers of  $[-N, N]$  together, one obtains the countable cover needed for LIND. For the second part, we only need to prove the forward implications. So assume LIND and note that the formula  $'(a_n, b_n) = I_x^\Psi'$  is just  $a_n =_{\mathbb{R}} x - \Psi(x) \wedge b_n =_{\mathbb{R}} x + \Psi(x)$ , which is  $\Pi_1^0$ , i.e. this formula is decidable using  $\mu^2$ , and we can treat it as quantifier-free in  $\text{ACA}_0^\omega$ . Now apply  $\text{QF-AC}^{0,1}$  to  $(\forall n \in \mathbb{N})(\exists x \in \mathbb{R})[(a_n, b_n) = I_x^\Psi]$  to obtain LIND<sub>2</sub>. For the final implication, we use the same argument as in the first part, establishing HBU relativised to  $[-N, N]$ , and now combined with the existence of the functional  $\Omega$  as in (3.1).  $\square$

We believe that the axiom  $\text{QF-AC}^{0,1}$  is essential in proving the second part. A similar observation is made by Kohlenbach in [32] regarding the local equivalence of 'epsilon-delta' and sequential continuity. We now have the following theorem.

**Theorem 3.14.** *The system  $\text{Z}_2^\omega$  proves LIND; no system  $\Pi_k^1\text{-CA}_0^\omega$  proves LIND.*

*Proof.* The second part is immediate from Theorems 3.1 and 3.13. The first part is proved by proving item (v) from Remark 1.1 in  $Z_2^\omega$ , and deriving LIND from this item. Thus, consider the following for any  $\Pi_\infty^1$ -formula  $A$  with any parameter:

$$(\forall f^1)(\exists n^0)A(\bar{f}n) \rightarrow (\exists \gamma^1 \in K_0)(\forall f^1)A(\bar{f}\gamma(f)). \quad (\text{NFP})$$

Here, ' $\gamma^1 \in K_0$ ' expresses that  $\gamma^1$  is an *associate*, which is the same as a *code* from RM by [33, Prop. 4.4]. Formally, ' $\gamma^1 \in K_0$ ' is the following formula:

$$(\forall f^1)(\exists n^0)(\gamma(\bar{f}n) >_0 0) \wedge (\forall n^0, m^0, f^1, ) (m > n \wedge \gamma(\bar{f}n) > 0 \rightarrow \gamma(\bar{f}n) =_0 \gamma(\bar{f}m)).$$

The value  $\gamma(f)$  for  $\gamma \in K_0$  is defined as the unique  $\gamma(\bar{f}n) - 1$  for  $n$  large enough. Now, since  $A$  as in NFP is a  $\Pi_k^1$ -formula for some  $k$ , we may treat it as quantifier-free given  $(\exists^3)$ . Applying QF-AC $^{1,0}$  to the antecedent of NFP, there is  $Y^2$  such that  $(\forall f^1)A(\bar{f}Y(f))$ . Define  $Z^2$  using  $(\exists^3)$  as follows:  $Z(f)$  is the least  $n \leq Y(f)$  such that  $A(\bar{f}n)$  if it exists, and zero otherwise. Note that  $Z$  is continuous on  $\mathbb{N}^{\mathbb{N}}$  and hence has an associate by [33, Prop. 4.7]. Alternatively, define the associate  $\gamma^1$  directly as follows: for  $w^{0*}$ , define  $\gamma(w)$  as the least  $n \leq |w|$  such that  $A(\bar{w}n)$  if such there is, and zero otherwise. Clearly, we have  $\gamma \in K_0$  and  $(\forall f^1)A(\bar{f}\gamma(f))$ , i.e. NFP follows. Finally, LIND follows from the latter by considering:

$$(\forall x \in \mathbb{R})(\exists n \in \mathbb{N})[(\exists y \in \mathbb{R})([x](\frac{1}{2^n}) - \frac{1}{n}, [x](\frac{1}{2^n}) + \frac{1}{n}) \subset I_y^\Psi] \quad (3.2)$$

for  $\Psi : \mathbb{R} \rightarrow \mathbb{R}^+$ , and where the formula in square brackets is abbreviated  $A(\bar{x}n)$ . This is a slight abuse of notation, as (only) the first  $2^n$  elements in the sequence  $x^1$  are being used in (3.2). Applying NFP to (3.2), we obtain  $\gamma \in K_0$  such that:

$$(\forall x \in \mathbb{R})(\exists y \in \mathbb{R})[[x](\frac{1}{2^{\gamma(x)}}) - \frac{1}{\gamma(x)}, [x](\frac{1}{2^{\gamma(x)}}) + \frac{1}{\gamma(x)}) \subset I_y^\Psi]. \quad (3.3)$$

Note that the formula in square brackets in (3.3) is arithmetical (including the formula needed to make the notation  $\gamma(x)$  work). Hence, using QF-AC and  $(\mu^2)$ , there is a functional  $\Phi$  which provides the real  $y$  from (3.3) on input  $x \in \mathbb{R}$ . The countable sub-cover of  $\cup_{x \in \mathbb{R}} I_x^\Psi$  can then be found by enumerating  $\Phi(q_w)$  for all finite sequences  $w^{0*}$  of rationals which represent rationals  $q_w^0$  and are such that  $\gamma(w) >_0 0$ . In particular, every  $x \in \mathbb{R}$  is in  $I_{\Phi(x)}^\Psi$  by (3.3), and since  $v^{0*} := \bar{x}2^{\gamma(x)}$  is in the aforementioned enumeration, we also have  $x \in I_{\Phi(q_v)}^\Psi$ .  $\square$

By the first part of Theorem 3.13, the results regarding LIND have to be somewhat similar to those for HBU. However, the Lindelöf theorem *for the Baire space* behaves quite differently, as will be established in Section 4.2. Furthermore, while HBU implies WKL, LIND does not by the following corollary.

**Corollary 3.15.** *The system  $\text{RCA}_0^\omega + \text{LIND}$  proves the same  $\text{L}_2$ -sentences as  $\text{RCA}_0$ .*

*Proof.* By the proof of [32, Prop. 3.1], if for a sentence  $A \in \text{L}_\omega$ , the system  $\text{RCA}_0^\omega$  proves  $A$ , then  $\text{RCA}_0$  proves  $[A]_{\text{ECF}}$ , where ' $[\cdot]_{\text{ECF}}$ ' is a syntactic translation which -intuitively- replaces any object of type 2 or higher by a code  $\gamma^1 \in K_0$ . Thus, to establish the corollary, it suffices to show that  $[\text{LIND}]_{\text{ECF}}$  is provable in  $\text{RCA}_0$ . However, LIND only involves objects of type 0 and 1, except for the leading quantifier. Hence,  $[\text{LIND}]_{\text{ECF}}$  is nothing more than LIND with ' $(\forall \Psi^{1 \rightarrow 1})$ ' replaced by ' $(\forall \gamma^1 \in K_0)$ '. Thus, by enumerating  $\gamma(w)$  as in the proof of the theorem, we immediately obtain a countable sub-cover, and  $[\text{LIND}]_{\text{ECF}}$  is provable in  $\text{RCA}_0$ .  $\square$

3.2.2. *Other theorems.* We discuss how the theorems in Remark 1.1 imply either LIND or HBU, and hence have similar properties to the latter.

- (1) The *Besicovitch and Vitali*<sup>5</sup> *covering lemmas* as in [1, §2] start from a cover of open balls, one for each  $x \in E \subset \mathbb{R}^n$ , and states the existence of a countable sub-cover of  $E$  with nice properties, i.e. LIND follows. Note that Vitali already (explicitly) discussed uncountable covers in [60, p. 236].
- (2) The existence of *Lebesgue numbers* for any open cover is equivalent to HBU, in the same way the countable case is equivalent to  $\text{WKL}_0$  ([24, Theorem 5.5]). The same holds for the *Banach-Alaoglu theorem*; the equivalence between the countable case and  $\text{WKL}_0$  is established in [10, p. 140].
- (3) The principle NFP implies LIND by the proof of Theorem 3.14.
- (4) The *Heine-Young* and *Lusin-Young* theorems from [62] are clearly refinements of HBU, while the *tile theorem* [30, 62], and the latter's generalisation due to Rademacher ([44, p. 190]) are clearly refinements of LIND.
- (5) Basic properties of the *gauge integral*, like uniqueness and its extension of the Riemann integral, are equivalent to HBU over the system  $\text{ACA}_0^\omega$ , as shown in Section 3.3. Note that  $\text{ACA}_0^\omega$  is very weak compared to  $\text{Z}_2^\omega$ , which is in turn required to prove HBU by Theorem 3.1.

Finally, we discuss how some of the ‘countable covering theorems’, like the Lindelöf and Vitali lemmas, from Remark 1.1 are used in mathematics.

**Remark 3.16.** The Cousin lemma is special because it deals with bounded sets (essentially the unit interval), while the other covering theorems apply to unbounded sets (e.g.  $\mathbb{R}^n$ ). Now, a cover of the latter is generally difficult to handle, but any *countable* sub-cover ‘automatically’ has nice properties: e.g. the *countable* sub-additivity of the Lebesgue measure. In fact, the proofs of *Sard's theorem* and the *maximal theorem* in [1], and of the *Lebesgue density theorem* in [54] are based on this idea. In other words, the non-local character of some of the covering theorems in Remark 1.1 has real applications in mathematics.

Similarly, for properties which hold in the unit interval *minus a measure zero set*, like the *differentiation theorem* for gauge integrals ([3, p. 80]), one uses the Vitali covering theorem to provide a countable sub-cover in which the complement of a finite sub-sub-cover has small length. Hence, one can neglect this complement and the finite nature of the sub-sub-cover then makes the proof straightforward.

**3.3. The gauge integral.** In this section, we provide a brief introduction to the *gauge integral* (Section 3.3.1) and establish that basic properties of this integral, like uniqueness and the fact it extends the Riemann and Lebesgue integral, are equivalent to HBU (Section 3.3.2) over the (relatively weak) system  $\text{ACA}_0^\omega$ . As will become clear below, the gauge integral enjoys both greater conceptual simplicity and higher generality than the Lebesgue integral. For this reason, there have been calls for (a somewhat stripped-down version of) the gauge integral to replace the Riemann and Lebesgue integral (and the associated measure theory) in the undergraduate curriculum ([4–6]). In a nutshell, the gauge integral can only be called *natural and mainstream*, and hence the same holds for HBU by the below equivalences.

---

<sup>5</sup>Not to be confused with the Vitali covering *theorem* ([3, p. 79]), which does follow from the Vitali covering *lemma* via Banach's proof from [40, p. 81]; we believe that the Vitali covering theorem is weaker than HBU, but nonetheless requires full second-order arithmetic  $\text{Z}_2^\omega$  to prove.



3.3.1. *Introducing the gauge integral.* The gauge integral is a generalisation of the Lebesgue and (improper) Riemann integral; it was introduced by Denjoy (in a different form) around 1912 and studied by Lusin, Perron, Henstock, and Kurzweil. The exact definition is in Definition 3.17, which we intuitively motivate as follows.

A limitation of the ‘ $\varepsilon$ - $\delta$ -definition’ of the Riemann integral is that near a singularity of a function  $f : [0, 1] \rightarrow \mathbb{R}$ , changes smaller than any fixed  $\delta > 0$  in  $x$  can still result in huge changes in  $f(x)$ , guaranteeing that the associated Riemann sums vary (much) more than the given  $\varepsilon > 0$ . The gauge integral solves this problem by replacing the *fixed*  $\delta > 0$  with a *gauge function*  $\delta : \mathbb{R} \rightarrow \mathbb{R}^+$ ; the latter can single out those partitions with ‘many’ partition points near singularities to compensate for the extreme behaviour there. Similarly,  $\delta : \mathbb{R} \rightarrow \mathbb{R}^+$  can single out partitions which avoid ‘small’ sets whose contribution to the Riemann sums should be negligible. We study  $\frac{1}{\sqrt{x}}$  and *Dirichlet’s function* in Example 3.18 after the following definition.

**Definition 3.17.** [Gauge integral]

- (1) A gauge on  $I \equiv [0, 1]$  is any function  $\delta : \mathbb{R} \rightarrow \mathbb{R}^+$ .
- (2) A sequence  $P := (t_0, I_0, \dots, t_k, I_k)$  is a *tagged partition* of  $I$ , written ‘ $P \in \mathbf{tp}$ ’, if the ‘tag’  $t_i \in \mathbb{R}$  is in the interval  $I_i$  for  $i \leq k$ , and the  $I_i$  partition  $I$ .
- (3) If  $\delta$  is a gauge on  $I$  and  $P = (t_i, I_i)_{i \leq k}$  is a tagged partition of  $I$ , then  $P$  is  $\delta$ -*fine* if  $I_i \subseteq [t_i - \delta(t_i), t_i + \delta(t_i)]$  for  $i \leq k$ .
- (4) For a tagged partition  $P = (t_i, I_i)_{i \leq k}$  of  $I$  and any  $f$ , the *Riemann sum*  $S(f, P)$  is  $\sum_{i=0}^n f(t_i)|I_i|$ , while the *mesh*  $\|P\|$  is  $\max_{i \leq n} |I_i|$ .
- (5) A function  $f : I \rightarrow \mathbb{R}$  is *Riemann integrable* on  $I$  if there is  $A \in \mathbb{R}$  such that  $(\forall \varepsilon >_{\mathbb{R}} 0)(\exists \delta >_{\mathbb{R}} 0)(\forall P \in \mathbf{tp})(\|P\| \leq_{\mathbb{R}} \delta \rightarrow |S(f, P) - A| <_{\mathbb{R}} \varepsilon)$ .
- (6) A function  $f : I \rightarrow \mathbb{R}$  is *gauge integrable* on  $I$  if there is  $A \in \mathbb{R}$  such that  $(\forall \varepsilon >_{\mathbb{R}} 0)(\exists \delta : \mathbb{R} \rightarrow \mathbb{R}^+)(\forall P \in \mathbf{tp})(P \text{ is } \delta\text{-fine} \rightarrow |S(f, P) - A| <_{\mathbb{R}} \varepsilon)$ .
- (7) A *gauge modulus* for  $f$  is a function  $\Phi : \mathbb{R} \rightarrow (\mathbb{R} \rightarrow \mathbb{R}^+)$  such that  $\Phi(\varepsilon)$  is a gauge as in the previous item for all  $\varepsilon >_{\mathbb{R}} 0$ .

The real  $A$  from items (5) and (6) in Definition 3.17 is resp. called the Riemann and gauge integral. We will always interpret  $\int_a^b f$  as a gauge integral, unless explicitly stated otherwise. We abbreviate ‘Riemann integration’ to ‘R-integration’, and the same for related notions. The following examples are well-known.

**Example 3.18** (Two examples). Let  $f$  be the function  $1/\sqrt{x}$  for  $x > 0$ , and zero otherwise. It is easy to show  $\int_0^1 f =_{\mathbb{R}} 2$  using the gauge modulus  $\delta_\varepsilon(x) := \varepsilon x^2$  for  $x > 0$  and  $\varepsilon^2$  otherwise. Let  $g$  be constant 1 for  $x \in \mathbb{Q}$ , and zero otherwise. It is easy to show  $\int_0^1 g =_{\mathbb{R}} 0$  using the gauge modulus  $\delta_\varepsilon(x) := 1$  if  $x \notin \mathbb{Q}$  and  $\varepsilon/2^{k+1}$  if  $x$  equals the  $k$ -th rational (for some enumeration of the rationals fixed in advance).

Finally, using the Axiom of Choice, a gauge integrable function always has a gauge modulus, but this is not the case in weak systems like  $\text{RCA}_0^\omega$ . However, to establish the *Cauchy criterion* for gauge integrals as in Theorem 3.21, a gauge modulus is essential. For this reason, we sometimes assume a gauge modulus when studying the RM of the gauge integral in Section 3.3.2. Similar ‘constructive enrichments’ exist in Friedman-Simpson RM, as established by Kohlenbach in [33, §4].

3.3.2. *Reverse Mathematics of the gauge integral.* We show that basic properties of the gauge integral are equivalent to HBU. We have based this development on Bartle’s introductory monograph [3].

First of all, we show that HBU is equivalent to the uniqueness of the gauge integral, and to the fact that the latter extends the R-integral. Note that the names of the two items in the theorem are from [3, p. 13-14].

**Theorem 3.19.** *Over  $\text{ACA}_0^\omega$ , the following are equivalent to HBU:*

- (i) *Uniqueness: If a function is gauge integrable on  $[0, 1]$ , then the gauge integral is unique.*
- (ii) *Consistency: If a function is R-integrable on  $[0, 1]$ , then it is gauge integrable there, and the two integrals are equal.*

*Proof.* We prove  $\text{HBU} \rightarrow \text{(i)} \rightarrow \text{(ii)} \rightarrow \text{HBU}$ , where only the first implication requires  $(\mu^2)$ . To prove that HBU implies *Uniqueness*, assume the former, let  $f$  be gauge integrable on  $I$  and suppose  $f$  satisfies for  $i = 1, 2$  (where  $A_i \in \mathbb{R}$ ) that:

$$(\forall \varepsilon > 0)(\exists \delta_i^1 : \mathbb{R} \rightarrow \mathbb{R}^+)(\forall P \in \mathbf{tp})(P \text{ is } \delta_i\text{-fine} \rightarrow |S(f, P) - A_i| < \varepsilon). \quad (3.4)$$

Fix  $\varepsilon > 0$  and the associated  $\delta_i : \mathbb{R} \rightarrow \mathbb{R}^+$  in (3.4) for  $i = 1, 2$ . We define the gauge  $\delta_3 : \mathbb{R} \rightarrow \mathbb{R}^+$  as  $\delta_3(x) := \min(\delta_1(x), \delta_2(x))$ . By definition, a partition which is  $\delta_3$ -fine, is also  $\delta_i$ -fine for  $i = 1, 2$ . Now assume there is  $P_0 \in \mathbf{tp}$  which is  $\delta_3$ -fine, and note that we obtain the following by applying (3.4) for the final inequality:

$$|A_1 - A_2| =_{\mathbb{R}} |A_1 - S(f, P_0) + S(f, P_0) - A_2| \leq_{\mathbb{R}} |A_1 - S(f, P_0)| + |S(f, P_0) - A_2| \leq_{\mathbb{R}} 2\varepsilon.$$

Hence, we must have  $A_1 =_{\mathbb{R}} A_2$ , and *Uniqueness* follows. What remains is to prove that for every gauge  $\delta$  there *exists* a  $\delta$ -fine tagged partition. We emphasise the crucial nature of this existence: (3.4) is vacuously true if there is no  $\delta_i$ -fine tagged partition; in other words: we can *only* make meaningful use of the conclusion of (3.4), *if* we show the existence of a  $\delta_i$ -fine tagged partition.

Thus, fix a gauge  $\delta : \mathbb{R} \rightarrow \mathbb{R}^+$  and apply HBU to  $\cup_{x \in I}(x - \delta(x), x + \delta(x))$  to obtain a finite sub-cover  $w := (y_0, \dots, y_k)$ , i.e. we have  $I \subset \cup_{x \in w}(x - \delta(x), x + \delta(x))$ . The latter cover is readily converted into a tagged partition  $P_0 := (z_j, I_j)_{j \leq l}$  (with  $l \leq k$  and  $z_j \in w$  for  $j \leq l$ ) by removing overlapping segments and omitting redundant intervals ‘from left to right’. By definition,  $z_j \in I_j \subset (z_j - \delta(z_j), z_j + \delta(z_j))$  for  $j \leq l$ , i.e.  $P_0$  is  $\delta$ -fine. While the previous two steps are straightforward, it should be noted that (i) HBU is essential by the equivalences in the theorem, and (ii) to convert  $w$  into a tagged partition, we need to compare real numbers (in the sense of deciding whether  $x >_{\mathbb{R}} 0$  or not) and this operation is only available in  $\text{ACA}_0^\omega$ .

To prove that *Uniqueness* implies *Consistency*, note that ‘ $P$  is  $d_\delta$ -fine’ is equivalent to ‘ $\|P\| \leq \delta$ ’ for the gauge  $d_\delta : \mathbb{R} \rightarrow \mathbb{R}^+$  which is constant  $\delta > 0$ . Rewriting the definition of Riemann integration with this equivalence, we observe that an R-integrable function  $f$  is also gauge integrable (with a constant gauge  $d_\delta$  for every choice of  $\varepsilon > 0$ ). The assumption *Uniqueness* then guarantees that the number  $A$  is the only possible gauge integral for  $f$  on  $I$ , i.e. the two integrals are equal.

To prove that *Consistency* implies HBU, suppose the latter is false, i.e. there is  $\Psi_0 : \mathbb{R} \rightarrow \mathbb{R}^+$  such that  $\cup_{x \in I} I_x^{\Psi_0}$  does not have a finite sub-cover. Now let  $f : I \rightarrow \mathbb{R}$  be R-integrable with R-integral  $A \in \mathbb{R}$ . Define the gauge  $\delta_0$  as  $\delta_0(x) := \Psi_0(x)$  and note that for any  $P \in \mathbf{tp}$ , we have that  $P$  is *not*  $\delta_0$ -fine, as  $\cup_{x \in I} I_x^{\Psi_0}$  would otherwise have a finite sub-cover (provided by the tags of  $P$ ). Hence, the following statement is vacuously true, as the underlined part is false:

$$(\forall \varepsilon > 0)(\forall P \in \mathbf{tp})(\underline{P \text{ is } \delta_0\text{-fine}} \rightarrow |S(f, P) - (A + 1)| < \varepsilon). \quad (3.5)$$

However, (3.5) implies that  $f$  is gauge integrable with gauge  $\delta_0$  and gauge integral  $A + 1$ , i.e. *Consistency* is false as the Riemann and gauge integrals of  $f$  differ. Note that  $\delta_0$  also provides a gauge *modulus* by (3.5) in case  $\neg\text{HBU}$ .  $\square$

In passing, we discuss the question if  $\text{ACA}_0^\omega$  in the previous (and subsequent) theorem can be weakened to  $\text{RCA}_0^\omega$ . In our opinion, this weakening would not be spectacular, given that HBU requires  $\mathbb{Z}_2^\omega$  for a proof, as established above. Furthermore, even very basic properties of the gauge integral require  $\text{ACA}_0^\omega$ , as follows.

**Example 3.20** (Splitting the domain). As it turns out, proving  $\int_0^1 f =_{\mathbb{R}} \int_0^x f + \int_x^1 f$  for  $0 <_{\mathbb{R}} x <_{\mathbb{R}} 1$  *in general* seems to require a *discontinuous* gauge. Indeed, if for  $\varepsilon > 0$  the functions  $\delta_1, \delta_2$  are gauges for the right-hand side of the equation, a gauge for the left-hand side is as follows ([3, p. 45]):

$$\delta_3(y) := \begin{cases} \min(\delta_1(y), \frac{1}{2}(x-y)) & y \in [0, x) \\ \min(\delta_1(x), \delta_2(x)) & y =_{\mathbb{R}} x \\ \min(\delta_2(y), \frac{1}{2}(y-x)) & y \in (x, 1] \end{cases} \quad (3.6)$$

The function  $\delta_3$  is discontinuous in general, but can be defined in  $\text{ACA}_0^\omega$ .

Secondly, we prove the *Cauchy criterion* for gauge integrals, as this theorem is needed below. Our proof is based on [3, p. 40] and illuminates the role of  $\Theta$ .

**Theorem 3.21** ( $\text{ACA}_0^\omega + \text{HBU}$ ; Cauchy criterion). *A function  $f : I \rightarrow \mathbb{R}$  is gauge integrable with a modulus if and only if there is  $\Phi : \mathbb{R}^+ \rightarrow (\mathbb{R} \rightarrow \mathbb{R}^+)$  such that*

$$(\forall \varepsilon >_{\mathbb{R}} 0)(\forall P, Q \in \text{tp})(P, Q \text{ are } \Phi(\varepsilon)\text{-fine} \rightarrow |S(f, P) - S(f, Q)| <_{\mathbb{R}} \varepsilon). \quad (3.7)$$

*Proof.* The forward implication follows by considering a gauge modulus  $\Phi$  for  $f$  and  $|S(f, P) - S(f, Q)| = |S(f, P) - A + A - S(f, Q)| \leq |S(f, P) - A| + |A - S(f, Q)| \leq \varepsilon$  where  $P, Q$  are  $\Phi(\varepsilon/2)$ -fine and  $A$  is the gauge integral of  $f$  over  $I$ . For the reverse implication let  $\Phi$  be as in (3.7); we need to find the real  $A$  from the definition of gauge integration. This real  $A$  can be obtained as the limit of the sequence  $S(f, Q_n)$  where  $Q_n$  is a  $\Phi(\frac{1}{2^n})$ -fine partition. Now, these partitions  $Q_n$  can in turn be defined by applying the functional  $\Omega$  from Theorem 3.3 to the canonical cover associated to  $\Phi(\frac{1}{2^n})$  and using Feferman's  $\mu$  to convert the resulting finite sub-cover to a suitable partition. Finally, (3.7) guarantees that the sequence  $S(f, Q_n)$  is Cauchy, while  $\text{ACA}_0$  proves that a Cauchy sequence has a limit by [55, III.2.2].  $\square$

The previous proof explains the need for a gauge modulus: the latter is essential in 'reconstructing' the gauge integral  $A$  as the limit in the proof, if  $A$  is not given.

Thirdly, we show that HBU is equivalent to the fact that the gauge integral encompasses the *improper* R-integral. The latter is a (usual) R-integral  $\int_a^b f(x)d(x)$  where additionally a limit operation like  $\lim_{a \rightarrow 0}$  or  $\lim_{b \rightarrow \infty}$  is applied. This method allows one to consider unbounded domains or use singularities as end points; as suggested by its name, an *improper* R-integral is (generally) not an actual R-integral. Now, *Hake's theorem* ([3, p. 195]) implies that improper R-integrals are *automatically* gauge integrals. We consider special cases of Hake's theorem, including item (iii) below which does mention gauge integrability *but does not mention gauge integrals or their uniqueness*. As a result, it is fair to say that the following equivalences are not (only) based on the uniqueness of the gauge integral.

**Theorem 3.22.** *Over  $\text{ACA}_0^\omega$ , the following are equivalent to HBU:*

- (i) *There exists a function which is not gauge integrable with a modulus.*
- (ii) *(Hake) If  $f$  is gauge integrable on  $I$  with a modulus and R-integrable on  $[x, 1]$  for  $x >_{\mathbb{R}} 0$ , then the limit of R-integrals  $\lim_{x \rightarrow 0+} \int_x^1 f$  is  $\int_0^1 f$ .*
- (iii) *(weak Hake) If  $f$  is gauge integrable with a modulus on  $I$  and R-integrable on  $[x, 1]$  for  $x >_{\mathbb{R}} 0$ , then the limit of R-integrals  $\lim_{x \rightarrow 0+} \int_x^1 f$  exists.*

*Proof.* We shall prove  $\text{HBU} \rightarrow (\text{ii}) \rightarrow (\text{iii}) \rightarrow (\text{i}) \rightarrow \text{HBU}$ . Note that the second implication is trivial. Now assume item (iii) and consider the function  $g : I \rightarrow \mathbb{R}$  which is 0 if  $x =_{\mathbb{R}} 0$ , and  $\frac{1}{x}$  otherwise. This function exists in  $\text{ACA}_0^\omega$  by [32, Prop. 3.12]. By the development of integration theory in [55, IV.2], the R-integral  $\int_x^1 g$  exists for  $x > 0$  and is readily seen to equal  $\ln(x)$ , the natural logarithm. However, the limit  $x \rightarrow 0+$  of this function is  $-\infty$ . Thus, the limit  $\lim_{x \rightarrow 0+} \int_x^1 g$  does not exist, and by the contraposition of weak Hake's theorem, we conclude that  $g$  is not gauge integrable with a modulus on  $I$ , i.e. item (i) follows.

The implication (i)  $\rightarrow$  HBU follows from the proof of Theorem 3.19: in the last part of the latter proof, it is shown that  $\neg\text{HBU}$  allows us to define a gauge  $\delta_0$  for which there are no  $\delta_0$ -fine partitions. Hence, the underlined part in (3.5) is false, making the formula trivially true for any  $f$  and  $A$ , i.e. every function is gauge integrable (with a modulus). Contraposition now yields the desired implication.

Finally, we prove item (ii) in  $\text{ACA}_0^\omega + \text{HBU}$  based on the proof in [3, p. 195]. In a nutshell, the latter uses the *Saks-Henstock lemma* to prove that the indefinite integral  $F(x) := \int_x^1 f$  is  $(\varepsilon\text{-}\delta)$ -continuous in  $x$  on  $I$ . Hence  $\lim_{x \rightarrow 0+} F(x) =_{\mathbb{R}} F(0)$ , which is exactly as required for item (ii). First of all, the Saks-Henstock lemma intuitively states that if one considers a sub-partition of a  $\delta$ -fine partition, one inherits all the 'nice' properties of the original partition. The proof of this lemma is a straight-forward 'epsilon-delta' argument, with one subtlety: the *Cauchy criterion* (as is Theorem 3.21) for gauge integrals requires a gauge modulus, which we (therefore) assumed in item (ii). The proof that the Saks-Henstock lemma yields the continuity of  $F(x) := \int_x^1 f$  is also a straight-forward 'epsilon-delta' argument.  $\square$

Fourth, we show that HBU is equivalent to the fact that the gauge integral is a proper extension of the Lebesgue integral. In fact,  $f : [0, 1] \rightarrow \mathbb{R}$  is Lebesgue integrable if and only if  $|f|, f$  are gauge integrable ([3, §7, p. 102]). We use the latter variant as introducing the Lebesgue integral is beyond the scope of this paper.

**Theorem 3.23.** *Over  $\text{ACA}_0^\omega$ , HBU is equivalent to the following statement: There exists a function  $\kappa : I \rightarrow \mathbb{R}$  which is gauge integrable with a modulus but  $|\kappa|$  is not.*

*Proof.* The reverse implication is immediate by Theorem 3.22. For the forward implication, define  $a_k := 1 - \frac{1}{2^k}$  and  $\kappa(x) := (-1)^{k+1} \frac{2^k}{k}$  if  $x \in [a_{k-1}, a_k)$  ( $k^0 \geq 1$ ), and 0 otherwise. Then for  $x >_{\mathbb{R}} 0$ , the area between the horizontal axis and the graph of  $|\kappa|$  on  $[0, x]$  is just a finite collection of (bounded) rectangles, i.e.  $|\kappa|$  is definitely R-integrable on  $[0, x]$  for  $x < 1$ . In particular, if  $x \geq_{\mathbb{R}} 1 - \frac{1}{2^k}$ , there are at least  $k$  rectangles to the left of  $x$ ; the first has base  $1/2$  and area 1, the second one base  $1/4$  and area  $1/2$ ,  $\dots$ , the  $k$ -th one has base  $1/2^k$  and area  $1/k$ . The R-integral  $\int_0^x |\kappa|$  is thus at least  $\sum_{i=1}^k \frac{1}{i}$ . The limit of the latter is the *divergent* harmonic series, and item (iii) from Theorem 3.22 yields that  $|\kappa|$  is not gauge integrable on  $I$  with a modulus. To prove that  $\kappa$  is gauge integrable on  $I$ , note that (3.6) allows us

to piece together gauges from sub-intervals. The following gauge modulus is based on that idea:

$$\delta_\varepsilon(x) := \begin{cases} d(x, E) & x \in [0, 1] \setminus E \\ \frac{\varepsilon}{4^{(k+1)}} & x =_{\mathbb{R}} a_k \\ 2^{-m(\varepsilon)} & x =_{\mathbb{R}} 1 \end{cases},$$

where  $E$  is the set consisting of the real 1 and all  $a_k$ , and where  $m(\varepsilon)$  is such that  $m(\varepsilon) \geq \frac{1}{\varepsilon}$  and the tail of the alternating harmonic series satisfies  $|\sum_{k=n}^{\infty} \frac{-1^{k+1}}{k}| \leq \varepsilon$  for  $n \geq m(\varepsilon)$ . We leave it as an exercise that this gauge can be defined in  $\text{ACA}_0^\omega$ . The proof that  $\delta_\varepsilon$  is a gauge for  $\kappa$  is completely straightforward and elementary, but somewhat long and tedious. Hence, we omit this proof and refer to [3, p. 35].  $\square$

Finally, we discuss in what sense we may evaluate general gauge integrals.

**Remark 3.24** (Computing integrals). In the case of the R-integral, a modulus (of R-integration) computes a  $\delta > 0$  in terms of any  $\varepsilon > 0$  as in Definition 3.17. Hence, if  $P_n$  is the equidistant partition of  $I$  with mesh  $1/2^n$ , we know that  $S(P_n, f)$  converges to the R-integral of  $f$  on  $I$ , and the modulus provides a rate of convergence. For the gauge integral, there is no analogue of the equidistant partition: even given a gauge modulus  $\delta(\varepsilon, x)$ , we need to find, say for every  $\varepsilon > 0$ , a  $\delta(\varepsilon, \cdot)$ -fine partition  $Q_\varepsilon$ ; only then can we consider the limit of  $S(Q_\varepsilon, f)$  for  $\varepsilon \rightarrow 0$ , which converges to the gauge integral of  $f$  on  $I$  as in Theorem 3.21. To find such a partition, the only option (we can imagine) is to consider  $\cup_{x \in I} (x - \delta(\varepsilon, x), x + \delta(\varepsilon, x))$  and apply the realiser  $\Omega$  for HBU as in (3.1) to obtain a finite sub-cover. The latter can be modified using  $\mu^2$  into a  $\delta(\varepsilon, \cdot)$ -fine partition.

#### 4. MAIN RESULTS II

We obtain some surprising results in higher-order RM related to the Cousin and Lindelöf lemmas. In particular, we study the behaviour of the latter in combination with Feferman's  $\mu$  and the Suslin functional  $S$ .

**4.1. Jumping to transfinite recursion.** In [42], we established that  $\text{ACA}_0^\omega + \text{HBU}$  proves arithmetical transfinite recursion as in  $\text{ATR}_0$ , i.e. HBU suffices to jump from the third to the fourth Big Five system. Moreover, we proved that there exists, without constructing it, a term of Gödel's  $T$  which computes a realiser for  $\text{ATR}_0$  from  $\Theta$  and  $\mu$ . Furthermore, in [43] we provide a direct construction of this fact, with a proof that  $\Theta$  *restricted to arithmetical functions suffices*. One consequence is that HBU cannot be satisfied within the class of Borel functions. Now, the space  $\bigoplus_{n \in \mathbb{N}} [0, 1]^n$  is a Polish space, and hence the following theorem makes perfect sense.

**Theorem 4.1.** *There is a map  $\Psi : [0, 1]^2 \rightarrow \mathbb{R}^+$  that is Borel at a finite level, and such that whenever  $\Phi : [0, 1] \rightarrow \bigoplus_{n \in \mathbb{N}} [0, 1]^n$  is Borel, there is an  $x \in [0, 1]$  such that  $\{I_y^{\Psi_x} \mid y \in \Phi(x)\}$  is not a covering of  $[0, 1]$ , where  $\Psi_x(y) = \Psi(x, y)$ .*

*Proof.* In [43, Cor. 4.14], we proved a similar result for Cantor space, and since the latter is homeomorphic to a nice closed subset of  $[0, 1]$ , the theorem follows.  $\square$

**4.2. Jumping to the superjump.** We show that the Lindelöf lemma *for Baire space* and Feferman's  $\mu^2$  together give rise to the Suslin functional  $S$  and the *superjump*  $\mathbb{S}$ . We introduce the latter in Section 4.2.1, while the Lindelöf lemma *for Baire space* and the associated functional  $\Xi$  (computing the countable sub-cover) are introduced in Section 4.2.2. The following results are established below.

- (1) The superjump  $\mathbb{S}$  is computable in the special fan functional  $\Theta$  and the Suslin functional  $S$  (Section 4.2.1).
- (2) The Suslin functional  $S$  is (uniformly) computable in Feferman's  $\mu$  and the functional  $\Xi$  which computes the countable sub-cover from the Lindelöf lemma for Baire space (Section 4.2.2).

As a consequence, the combination of Feferman's  $\mu$  and any such  $\Xi$  computes the superjump  $\mathbb{S}$ . We recall the fact that the special fan functional  $\Theta$  is not unique, and neither is 'the' aforementioned functional  $\Xi$ .

4.2.1. *Computing the superjump.* We show that the combination of the Suslin functional  $S$  and the special fan functional  $\Theta$  computes the *superjump*. The latter corresponds to the Halting problem for computability on type two inputs. Indeed, the superjump  $\mathbb{S}^3$  was introduced in [23] by Gandy (essentially) as follows:

$$\mathbb{S}(F^2, e^0) := \begin{cases} 0 & \text{if } \{e\}(F) \text{ terminates} \\ 1 & \text{otherwise} \end{cases}, \quad (\mathbb{S}^3)$$

where the formula ' $\{e\}(F)$  terminates' is a  $\Pi_1^1$ -formula defined by Kleene's S1-S9.

As to its history, Harrington has proved that the first ordinal not computable in  $\mathbb{S}$  is the first recursively Mahlo ordinal ([25]). In turn, the latter ordinal appears in the study of constructive set and type theory and the associated proof theory ([45–47]). In particular,  $\{R \subseteq \mathbb{N} : R \text{ is computable from } \mathbb{S}\}$  is the smallest  $\beta$ -model of  $\Delta_2^1\text{-CA}_0 + (\mathbf{M})$ , where  $(\mathbf{M})$  expresses that every true  $\Pi_3^1$ -sentence with parameters already holds in a  $\beta$ -model of  $\Delta_2^1$ -comprehension ([47]). As discussed in Remark 4.15,  $\mathbb{S}$  lives far outside of predicative mathematics.

**Theorem 4.2.** *The superjump  $\mathbb{S}$  is computable in any  $\Theta$  satisfying  $\text{SCF}(\Theta)$  and the Suslin functional  $S$ .*

*Proof.* We first provide a sketch of the proof as follows. Recall that if  $\sigma$  is a finite binary sequence, then  $[\sigma]$  is the set of total binary extensions of  $\sigma$ .

- (1) Given  $F^2$ , let  $\alpha_F(e) = \{e\}(F, e)$  whenever the value is in  $\{0, 1\}$ , and let  $X_F$  be the set of total binary extensions of  $\alpha_F$ .
- (2) Compute  $G_F$  from  $F$  and  $S$  with the properties
  - i) if  $f \notin X$ , then  $G_F(f) > 0$
  - ii) if  $G_F(f) > 0$ , then  $[\bar{f}G_F(f)]$  does not intersect  $X_F$
  - iii) for  $f \in X_F$ ,  $G_F(f) = 0$ .
- (3) Show that  $\mathbb{S}(F)$  is uniformly computable in  $S$  and  $f \in X_F$ .
- (4) Since  $\Theta(G_F)$  has to intersect  $X_F$ , and we can decide where,  $\mathbb{S}(F)$  is computable in  $\Theta$  and  $S$ , uniformly in  $F$ .

We work out the proof in full detail below. □

We will now list some basic lemmas needed for the detailed proof of Theorem 4.2. We first define an important concept relating to S1-S9 computability with type two inputs. Its importance stems from the fact that it is independent of the choice of input functional  $F^2$ , as follows.

**Lemma 4.3.** *There is a primitive recursive  $\xi$  of type level 1, independent of the choice of  $F^2$ , such that  $\{\xi(e, \vec{a})\}(F, \xi(e, \vec{a}))$  is resp.  $(0, 1, \text{undefined})$  whenever  $\{e\}(F, \vec{a})$  is resp.  $(= 0, > 0, \text{undefined})$ .*

**Lemma 4.4.** *There is a primitive recursive function  $\eta$  such that for all  $e, \vec{a}, F$*

$$\{\eta(e)\}(F, \vec{a}) \simeq \{e\}(F, \vec{a}) \dot{-} 1,$$

where ‘ $\simeq$ ’ means that both sides are undefined or both sides are defined and equal.

**Definition 4.5.** Let  $f$  be a total binary function. By an application of the recursion theorem for Turing computations in oracles we define

$$[e]_f(\vec{a}) := \begin{cases} 0 & \text{if } f(\xi(e, \vec{a})) = 0 \\ 1 + [\eta(e)]_f(\vec{a}) & \text{if } f(\xi(e, \vec{a})) = 1 \end{cases}.$$

Clearly, if the recursion goes on forever,  $[e]_f(\vec{a})$  will be undefined.

Intuitively speaking and from the outside,  $[\cdot]_f$  may look like an indexing of some partial functions computable in some functional of type 2, but to what extent this is correct, will depend on the choice of  $f$ .

We will now use  $F$  to define a relation, mimicking the subcomputation relation relative to  $F$ , as far as possible. As a cheap trick, we will let an alleged computation tuple be a subcomputation of its own if it is clear that something is wrong, in order to force such objects into the non-well-founded part of the relation.

**Definition 4.6.** Given  $f$ , we let  $\Omega_f$  be the set of triples  $(e, \vec{a}, b)$  such that  $[e]_f(\vec{a}) = b$ . Given  $F$  as well, define the relation ‘ $\prec$ ’ (short for  $\prec_{f,F}$ ) on  $\Omega_f$  as follows:

- If  $e$  is not a Kleene index for any of S1-S9, we put  $(e, \vec{a}, b) \prec (e, \vec{a}, b)$ .
- If  $e$  is an index for an initial computation, we let  $(e, \vec{a}, b)$  be a leaf in our ordering if  $\{e\}(F, \vec{a}) = b$ , and its own sub-node otherwise. This decision will be independent of the choice of the functional  $F$ .
- We treat the case S4. The rest of the cases, except S8, are similar or void (e.g. S6). If  $e$  is an index for composition  $\{e\}(F, \vec{a}) = \{e_1\}(\{e_2\}(\vec{a}), \vec{a})$ ,  $c$  is given and there is a  $b$  such that  $[e_2]_f(\vec{a}) = b$ ,  $[e_1]_f(b, \vec{a}) = c$  and  $[e]_f(\vec{a}) = c$ , then we define  $(e_2, \vec{a}, b) \prec (e, \vec{a}, c)$  and  $(e_1, b, \vec{a}, c) \prec (e, \vec{a}, c)$ . If there is no such  $b$ , we let  $(e, \vec{a}, c) \prec (e, \vec{a}, c)$ .
- For the case S8, if we have  $\{e\}(F, \vec{a}) = F(\lambda b. \{d\}(F, b, \vec{a}))$ , we let  $(e, \vec{a}, c) \prec (e, \vec{a}, c)$  unless  $h(b) = [d]_f(b, \vec{a})$  is a total function and  $F(h) = c$ . In the latter case, we let  $(d, b, \vec{a}, h(b)) \prec (e, \vec{a}, c)$  for all  $b$ .

The intuitive explanation of Definition 4.6 is as follows: The set of finite sequences  $(e, \vec{a}, b)$  such that  $\{e\}(F, \vec{a}) = b$  is defined by a strictly positive inductive definition, so whenever a sequence is added to the set it is either initial or there is a unique set of other sequences in the set causing that we accept the one chosen. These are called *immediate* predecessors in the computation tree. The relation ‘ $\prec$ ’ is defined on the set of  $(e, \vec{a}, b)$  where  $[e]_f(\vec{a}) = b$  as the immediate predecessor relation wherever the inductive definition of the computation tree is locally correct.

**Lemma 4.7.** *For any function  $f$ , the well-founded segment of  $\langle \Omega_f, \prec_{f,F} \rangle$  is an initial segment of the full computation relation of  $F$ .*

*Proof.* This is trivial by induction over this well-founded segment.  $\square$

**Lemma 4.8.** *For any  $f \in X_F$ , if  $\{e\}(F, \vec{a}) = b$ , then  $[e]_f(\vec{a}) = b$ .*

*Proof.* We prove this by induction on  $b$ . If  $b = 0$ , then  $\{\xi(e, \vec{a})\}(F, \xi(e, \vec{a})) = 0$ , so  $f(\xi(e, \vec{a})) = 0 = [e]_f(\vec{a})$ . If  $b > 0$ , we use the induction hypothesis on  $b \dot{-} 1$  for the index  $\eta(e)$  and the fact that  $[e]_f(\vec{a}) = b$  in this case.  $\square$

**Lemma 4.9.** *If  $f \in X_F$  and  $\{e\}(F, \vec{a}) = b$ , then  $(e, \vec{a}, b)$  is in the  $\prec_{f,F}$ -well-founded part of  $\Omega_f$ . Moreover, this well-founded part is exactly the full tree of terminating computations  $\{e\}(F, \vec{a}) = b$  relative to  $F$ .*

*Proof.* That the computation tree for computations relative to  $F$  is contained in the well-founded part is proved by induction over the tree of real computations. Now, if the well-founded part of  $\langle \Omega_f, \prec_{f,F} \rangle$  contains more, we may consider one alleged computation  $(e, \vec{a}, b)$  in  $\Omega_f$  that is not a real  $F$ -computation, but that is minimal as such. Since it is in the well-founded part,  $(e, \vec{a}, b)$  is locally correct, so either it is an initial computation or it has subcomputations that are real (because we consider a minimal one). Being locally correct, we see in each case that  $(e, \vec{a}, b)$  must be genuine after all.  $\square$

**Lemma 4.10.** *If  $f \in X_F$ , then  $\mathbb{S}(F)$  is uniformly computable in  $f$ ,  $F$  and  $S$ .*

*Proof.* From the data, we can compute the characteristic function of  $\{(e, \vec{a}, b) \mid \{e\}(\vec{a}) = b\}$ , and  $\mathbb{S}(F)$  is primitive recursive in this characteristic function.  $\square$

We are now ready to provide the proof of Theorem 4.2 as follows.

*Proof.* We see from Lemma 4.7 that if the  $\prec_{f,F}$ -well-founded part of  $\Omega_f$  is closed under the Kleene schemes S1-S9 relative to  $F$ , then  $\mathbb{S}(F)$  is computable in  $f$ ,  $F$  and  $S$  as above. We need  $S$  to isolate the well-founded part, and (only)  $F$  and  $\mu^2$  to decide if we have the closure.

Now, assume that  $f$  is such that the  $\prec_{f,F}$ -well-founded part is not S1-S9-closed. Let  $\{e\}(F, \vec{a}) = b$  be a computation of minimal rank such that we do not have  $[e]_f(\vec{a}) = b$ . By induction on  $b$  we see that there must be an index  $d$  such that  $\{d\}(F, d) \in \{0, 1\}$  and  $\{d\}(F, d) \neq f(d)$ . If we then put  $G_F(f) := d + 1$  we have ensured that there will be no extension of  $\bar{f}G_F(f)$  in  $X_F$ . Using Gandy selection for  $F, \mu$  and  $f$ , we can trivially find a  $d$  with this property from the well-founded part of  $\Omega_f$ . In order to show that  $G_F$  is definable from  $S, F, \mu$  via a term in Gödel's  $T$ , we proceed as follows:

Given the well-founded part  $W$  of  $\Omega_F$ , we may arithmetically decide if it respects S1-S9. If it does not, let  $\Gamma$  be the, arithmetically in  $F$ , inductive definition of the computation tuples for computing relative to  $F$ , and by one application of  $\mu$  on  $\Gamma(W) \setminus W$ , we may find the  $(e, \vec{a}, b)$  that leads us to the  $d$  we need.

In light of the previous, we put  $G_F(f) := 0$  if the  $\prec_{f,F}$ -well-founded part of  $\Omega_f$  is a fixed point of the inductive definition of computations relative to  $F$ , while we put  $G_F(f) := d + 1$  for the  $d$  selected as above otherwise. Thus,  $\Theta(G_F)$  must contain a function from which, together with  $F$  and  $S$ , we can compute  $\mathbb{S}(F)$ .  $\square$

**4.2.2. Computing the Suslin functional.** We show that the Suslin functional  $S$  can be computed by the combination of Feferman's  $\mu$  and the functional  $\Xi$  arising from the Lindelöf lemma for  $\mathbb{N}^{\mathbb{N}}$ . Regarding the latter, we recall that Lindelöf already proved that Euclidean space is *hereditarily Lindelöf* in [34] around 1903. Now, the latter hereditary property implies that  $\mathbb{N}^{\mathbb{N}}$  has the Lindelöf property, since  $\mathbb{N}^{\mathbb{N}}$  is homeomorphic to the irrationals in  $[0, 1]$  using continued fractions expansion.

Now, for any  $\Psi^2$ , the corresponding ‘canonical cover’ of  $\mathbb{N}^{\mathbb{N}}$  is  $\cup_{f \in \mathbb{N}^{\mathbb{N}}} [\bar{f}\Psi(f)]$  where  $[\sigma^{0^*}]$  is the set of all extensions in  $\mathbb{N}^{\mathbb{N}}$  of  $\sigma$ . By the Lindelöf lemma for  $\mathbb{N}^{\mathbb{N}}$ ,



there is a sequence  $f_{(\cdot)}^{0 \rightarrow 1}$  such that the set of  $\cup_{i \in \mathbb{N}} [\bar{f}_i \Psi(f_i)]$  still covers  $\mathbb{N}^{\mathbb{N}}$ , i.e.

$$(\forall \Psi^2)(\exists f_{(\cdot)}^{0 \rightarrow 1})(\forall g^1)(\exists n^0)(g \in [\bar{f}_n \Psi(f_n)]). \quad (\text{LIND}_4)$$

Similar to the specification  $\text{SCF}(\Theta)$  for the special fan functional  $\Theta$ , we introduce the following specification based on  $\text{LIND}_4$ . As for the former specification, the functional  $\Xi^{2 \rightarrow (0 \rightarrow 1)}$  satisfying  $\text{LIN}(\Xi)$  is not unique.

$$(\forall \Psi^2)(\forall g^1)(\exists n^0)(g \in [\overline{\Xi(\Psi)(n)} \Psi(\Xi(\Psi)(n))]). \quad (\text{LIN}(\Xi))$$

As for the special fan functional  $\Theta$  in Theorem 3.3, the existence of  $\Xi$  as in  $\text{LIN}(\Xi)$  amounts to the Lindelöf lemma  $\text{LIND}_4$  itself.

**Theorem 4.11.** *The system  $\Pi_1^1\text{-CA}_0^\omega$  proves  $\text{LIND}_4 \leftrightarrow (\exists \Xi)\text{LIN}(\Xi)$ .*

*Proof.* We only need to prove the forward direction. We rephrase  $\text{LIND}_4$  to

$$(\forall G^2)(\exists f_{(\cdot)}^{0 \rightarrow 1})[(\forall g^1)(\exists n^0)(g \in [\bar{f}_n^+ f_n(0)]) \wedge (\forall m^0)(f_m(0) = G(f_m^+))], \quad (4.1)$$

where  $f^+(k) = f(k+1)$ . Using the Suslin functional  $S$  and  $\mu$  we see that the part of (4.1) inside the (outermost) square brackets can be viewed as quantifier-free, and thus the existence of  $\Xi$  follows from  $\text{QF-AC}$ .  $\square$

**Theorem 4.12.**  *$\text{RCA}_0^\omega + (\exists \Xi)\text{LIN}(\Xi)$  proves the same  $\text{L}_2$ -sentences as  $\text{RCA}_0$ .*

*Proof.* As in the proof of Corollary 3.15, it suffices to show that  $[(\exists \Xi)\text{LIN}(\Xi)]_{\text{ECF}}$  is provable in  $\text{RCA}_0$ . However,  $(\exists \Xi)\text{LIN}(\Xi)$  only involves objects of type 0 and 1 except for the two leading quantifiers. Hence,  $[(\exists \Xi)\text{LIN}(\Xi)]_{\text{ECF}}$  is as follows:

$$(\exists \xi^1 \in K_0)(\forall \gamma^1 \in K_0)(\forall g^1)(\exists n^0)(g \in [\overline{\xi(\gamma)(n)} \gamma(\xi(\gamma)(n))]).$$

Thus, by defining  $\xi$  as the enumeration of  $\gamma(w)$  as in the proof of Theorem 3.14, we obtain an associate for a functional producing a countable sub-cover, and the sentence  $[(\exists \Xi)\text{LIN}(\Xi)]_{\text{ECF}}$  is therefore provable in  $\text{RCA}_0$ .  $\square$

**Theorem 4.13.** *The Suslin functional  $S$  is uniformly computable in Feferman's  $\mu$  and any  $\Xi$  satisfying  $\text{LIN}(\Xi)$ . Furthermore,  $\text{ACA}_0^\omega + (\exists \Xi)\text{LIN}(\Xi)$  proves  $(S^2)$*

*Proof.* Recall the definition of the Suslin functional  $S$  as follows:

$$S(f) = \begin{cases} 0 & \text{if } (\exists g^1)(\forall n^0)(f(\bar{g}n) = 0) \\ 1 & \text{otherwise} \end{cases}.$$

Define  $F_f^2(g)$  as  $n+1$  if  $n$  is minimal such that  $f(\bar{g}n) > 0$ , and 0 if there is no such  $n$ . Note that  $F_f$  is readily defined from  $f$  (in terms of  $\mu^2$ ) inside  $\text{ACA}_0^\omega$ , and note that if  $F_f(h) > 0$  and  $\bar{g}F_f(h) = \bar{h}F_f(h)$ , then  $F_f(g) = F_f(h)$ . Let  $\Xi$  be such that  $\text{LIN}(\Xi)$ , and consider the following formula

$$S(f) = 0 \leftrightarrow (\exists i^0)(F_f(\Xi(F_f)(i)) = 0). \quad (4.2)$$

The reverse direction in (4.2) is immediate by the definition of  $F_f$ . For the forward direction, assume  $S(f) = 0$  and let  $g^1$  satisfy  $(\forall n^0)(f(\bar{g}n) = 0)$ , i.e.  $F_f(g) = 0$ . As observed above, if  $F_f(h) > 0$ , we have  $g \notin [\bar{h}F_f(h)]$ ; hence if  $F_f(h_n) > 0$  for all  $n \in \mathbb{N}$  where  $h_n = \Xi(F_f)(n)$ , the corresponding countable subset of the covering induced by  $F_f$  will not be a covering. Thus  $F_f(\Xi(F_f)(n)) = 0$  must hold for some  $n$ , i.e. the right-hand side of (4.2) follows. Finally, (4.2) clearly characterises  $S(f)$  in terms of  $\mu$ ,  $f$  and  $\Xi$  (via a term in Gödel's  $T$ ), and we are done.  $\square$

Combining the above results, we obtain the following theorem.

**Theorem 4.14.** *The superjump  $\mathbb{S}$  is computable in any  $\Xi$  satisfying  $\text{LIN}(\Xi)$  and Feferman's  $\mu$ , by a term in Gödel's  $T$ .*

*Proof.* Given such  $\Xi$ , there are terms  $t_1, t_2$  such that  $\text{SCF}(t_1(\Xi, \mu))$  (i.e.  $\Theta$  is given by  $t_1(\Xi, \mu)$ ), and  $S =_2 t_2(\Xi, \mu)$ . Checking the details of the proof of Theorem 4.2 and the construction of  $G_F$ , we see that there is a term  $t_3$  such that  $G_F(f) = t_3(F, f, S, \mu)$ . Since  $\mathbb{S}(F)$  is primitive recursive in  $\Theta(G_F)$ , the theorem follows.  $\square$

**Remark 4.15** (On predicativist mathematics). We have discussed the *compatibility problem* for Nelson's predicative arithmetic (and its negative answer) in Section 1.2.3. We now argue that Theorems 4.12 and 4.13 settle the compatibility problem for Weyl-Feferman predicative mathematics *in the negative*. To this end, we exhibit two theorems  $A$  and  $B$  which are both *acceptable* in predicative mathematics but  $A \wedge B$  is not. In a nutshell,  $\text{ATR}_0$  is considered the 'upper limit' of predicative mathematics; both  $\text{RCA}_0^\omega + (\exists\Xi)\text{LIN}(\Xi)$  and  $\text{ACA}_0^\omega$  fall 'well below' this upper limit, while the combination  $\text{ACA}_0^\omega + (\exists\Xi)\text{LIN}(X)$  falls 'well above' the upper limit. Hence, each of the former systems is acceptable in predicative mathematics, but the combination is not. A more detailed discussion, for those familiar with the required technical notions from mathematical logic, is as follows:

First of all, we elaborate on the notion of 'acceptable in predicative mathematics'. On one hand there is Feferman's notion of *predicative provability* ([17, 20]), which is rather limited and clumsy when dealing with ordinary mathematics, according to Simpson ([56, p. 154]). On the other hand, the weaker notion of *predicative reducibility* is more flexible: a formal system  $T$  is *predicatively reducible* if -intuitively speaking- it is not stronger than a system  $S$  which is predicatively provable. Thus, while  $T$  may involve *impredicative* notions, the latter are 'safe' from the point of view of predicative mathematics as these notions only provide as much strength/power as  $S$ , and the latter's 'predicative status' is well-known.

Secondly, Feferman and Schütte have shown (independently) that the least *non-predicatively provable* ordinal is  $\Gamma_0$  (See [20, p. 607] for details and references). Hence, a formal system  $T$  is called *predicatively reducible* if its ordinal  $|T|$  satisfies  $|T| < \Gamma_0$ . Note that  $|\text{ATR}_0| = \Gamma_0$ , which motivates the status of  $\text{ATR}_0$  as the upper limit of predicative mathematics. Now, the proof-theoretic ordinal of  $\text{RCA}_0^\omega + (\exists\Xi)\text{LIN}(\Xi)$  (resp.  $\text{ACA}_0^\omega$ ) is  $\omega^\omega$  (resp.  $\varepsilon_0$ ) by Theorem 4.12 (resp. [48, Theorem 2.2]) and [55, IX.5]. Since  $\omega^\omega < \varepsilon_0 < \Gamma_0$ , both these systems are predicatively reducible. By contrast, the combination of these systems, namely  $\text{ACA}_0^\omega + (\exists\Xi)\text{LIN}(\Xi)$  implies  $\Pi_1^1\text{-CA}_0$  by Theorem 4.13, and the ordinal for the latter system is far above  $\Gamma_0$ . We refer to [55, IX.5] for background concerning the cited results and further references.

**Acknowledgement 4.16.** Our research was supported by the John Templeton Foundation, the Alexander von Humboldt Foundation, LMU Munich (via the Excellence Initiative and the Center for Advanced Studies of LMU), and the University of Oslo. The authors express their gratitude towards these institutions. The authors thank Ulrich Kohlenbach and Anil Nerode for their valuable advice.

## REFERENCES

- [1] Pascal Auscher and Lashi Bandara, *Real Harmonic Analysis*, ANU Press, 2010.
- [2] Jeremy Avigad and Solomon Feferman, *Gödel's functional ("Dialectica") interpretation*, Handbook of proof theory, Stud. Logic Found. Math., vol. 137, 1998, pp. 337–405.

- [3] Robert G. Bartle, *A modern theory of integration*, Graduate Studies in Mathematics, vol. 32, American Mathematical Society, 2001.
- [4] ———, *Return to the Riemann integral*, Amer. Math. Monthly **103** (1996), no. 8.
- [5] Robert G. Bartle, Ralph Henstock, Jaroslav Kurzweil, Eric Schechter, Stefan Schwabik, and Rudolf Vyborný, *An open letter*, Website: <https://math.vanderbilt.edu/schectex/ccc/gauge/letter/> (1997).
- [6] Robert G. Bartle, *Book Review: The general theory of integration*, Bull. Amer. Math. Soc. (N.S.) **29** (1993), no. 1, 136–139.
- [7] Jon Barwise (Ed.), *Handbook of mathematical logic*, North-Holland, 1977. Studies in Logic and the Foundations of Mathematics, Vol. 90.
- [8] Nicolas Bourbaki, *General topology. Chapters 1–4*, Springer, 1998.
- [9] Emile Borel, *Sur quelques points de la théorie des fonctions*, Ann. Sci. École Norm. Sup. (3) **12** (1895), 9–55.
- [10] Douglas K. Brown, *Functional analysis in weak subsystems of second-order arithmetic*, PhD Thesis, The Pennsylvania State University, ProQuest LLC, 1987.
- [11] Frank E. Burk, *A garden of integrals*, The Dolciani Mathematical Expositions, vol. 31, Mathematical Association of America, Washington, DC, 2007.
- [12] Samuel R. Buss, *Nelson’s work on logic and foundations and other reflections on the foundations of mathematics*, Diffusion, quantum theory, and radically elementary mathematics, Math. Notes, vol. 47, Princeton Univ. Press, Princeton, NJ, 2006, pp. 183–208.
- [13] Pierre Cousin, *Sur les fonctions de  $n$  variables complexes*, Acta Math. **19** (1895), 1–61.
- [14] L. P. G. Dirichlet, *Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données* (2008). <https://arxiv.org/abs/0806.1294>.
- [15] Pierre Dugac, *Sur la correspondance de Borel et le théorème de Dirichlet-Heine-Weierstrass-Borel-Schoenflies-Lebesgue*, Arch. Internat. Hist. Sci. **39** (1989), no. 122, 69–110.
- [16] Damir D. Dzhabarov, *Reverse Mathematics Zoo*. <http://rmzoo.uconn.edu/>.
- [17] Solomon Feferman, *Systems of predicative analysis*, J. Symbolic Logic **29** (1964), 1–30.
- [18] ———, *In the light of logic*, Logic and Computation in Philosophy, OUP, 1998.
- [19] ———, *The significance of Weyl’s Das Kontinuum*, Proof theory (Roskilde, 1997), Synthese Lib., vol. 292, Kluwer Acad. Publ., Dordrecht, 2000, pp. 179–194.
- [20] ———, *Predicativity*, The Oxford Handbook of the Philosophy of Mathematics and Logic, Oxford University Press, 2005, pp. 590–624.
- [21] Harvey Friedman, *Some systems of second order arithmetic and their use*, Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 1, 1975, pp. 235–242.
- [22] ———, *Systems of second order arithmetic with restricted induction, I & II (Abstracts)*, Journal of Symbolic Logic **41** (1976), 557–559.
- [23] R. O. Gandy, *General recursive functionals of finite type and hierarchies of functions*, Ann. Fac. Sci. Univ. Clermont-Ferrand No. **35** (1967), 5–24.
- [24] Mariagnese Giusto and Alberto Marcone, *Lebesgue numbers and Atsugi spaces in subsystems of second-order arithmetic*, Arch. Math. Logic **37** (1998), no. 5-6, 343–362.
- [25] Leo Harrington, *The superjump and the first recursively Mahlo ordinal*, Generalized recursion theory (Proc. Sympos., Univ. Oslo, Oslo, 1972), North-Holland, 1974, pp. 43–52. Studies in Logic and the Foundations of Math., Vol. 79.
- [26] John P. Hartley, *Effective discontinuity and a characterisation of the superjump*, J. Symbolic Logic **50** (1985), no. 2, 349–358.
- [27] David Hilbert, *David Hilbert’s lectures on the foundations of arithmetic and logic, 1917–1933*, Springer, 2013. Edited by William Ewald, Wilfried Sieg and Michael Hallett.
- [28] David Hilbert and Paul Bernays, *Grundlagen der Mathematik. I*, Zweite Auflage. Die Grundlehren der mathematischen Wissenschaften, Band 40, Springer, 1968.
- [29] ———, *Grundlagen der Mathematik. II*, Zweite Auflage. Die Grundlehren der mathematischen Wissenschaften, Band 50, Springer, 1970.
- [30] T. H. Hildebrandt, *The Borel theorem and its generalizations*, Bull. Amer. Math. Soc. **32** (1926), no. 5, 423–474.
- [31] Israel Kleiner, *Excursions in the history of mathematics*, Birkhäuser/Springer, 2012.
- [32] Ulrich Kohlenbach, *Higher order reverse mathematics*, Reverse mathematics 2001, Lect. Notes Log., vol. 21, ASL, 2005, pp. 281–295.
- [33] ———, *Foundational and mathematical uses of higher types*, Reflections on the foundations of mathematics (Stanford, CA, 1998), Lect. Notes Log., vol. 15, ASL, 2002, pp. 92–116.

- [34] Ernst Lindelöf, *Sur Quelques Points De La Théorie Des Ensembles.*, Comptes Rendus (1903), 697–700.
- [35] John Longley and Dag Normann, *Higher-order Computability*, Theory and Applications of Computability, Springer, 2015.
- [36] Colin McLarty, *Interpreting set theory in higher order arithmetic*, arXiv: <https://arxiv.org/abs/1207.6357> **72** (2013, v5), pp. 6.
- [37] Antonio Montalbán, *Open questions in reverse mathematics*, Bull. Symbolic Logic **17** (2011), no. 3, 431–454.
- [38] P. Muldowney, *A general theory of integration in function spaces, including Wiener and Feynman integration*, Vol. 153, Longman Scientific & Technical, Harlow; John Wiley, 1987.
- [39] Carl Mummert and Stephen G. Simpson, *Reverse mathematics and  $\Pi_2^1$  comprehension*, Bull. Symbolic Logic **11** (2005), no. 4, 526–533.
- [40] I. P. Natanson, *Theory of functions of a real variable*, Frederick Ungar Publishing Co., 1955.
- [41] Edward Nelson, *Predicative arithmetic*, Mathematical Notes, vol. 32, Princeton University Press, Princeton, NJ, 1986.
- [42] Dag Normann and Sam Sanders, *Nonstandard Analysis, Computability Theory, and their connections*, Submitted; arXiv: <https://arxiv.org/abs/1702.06556> (2017).
- [43] ———, *Nonstandard Analysis, Computability Theory, and metastability*, In Preparation (2017).
- [44] Hans Rademacher, *Eineindeutige Abbildungen und Meßbarkeit*, Monatsh. Math. Phys. **27** (1916), no. 1, 183–235.
- [45] Michael Rathjen, *The superjump in Martin-Löf type theory*, Logic Colloquium '98 (Prague), Lect. Notes Log., vol. 13, Assoc. Symbol. Logic, Urbana, IL, 2000.
- [46] ———, *Ordinal notations based on a weakly Mahlo cardinal*, Arch. Math. Logic **29** (1990), no. 4, 249–263.
- [47] ———, *The recursively Mahlo property in second order arithmetic*, Math. Logic Quart. **42** (1996), no. 1, 59–66.
- [48] Nobuyuki Sakamoto and Takeshi Yamazaki, *Uniform versions of some axioms of second order arithmetic*, MLQ Math. Log. Q. **50** (2004), no. 6, 587–593.
- [49] Gerald E. Sacks, *Higher recursion theory*, Perspectives in Mathematical Logic, Springer, 1990.
- [50] Sam Sanders, *The Gandy-Hyland functional and a computational aspect of Nonstandard Analysis*, To appear in *Computability*, arXiv: <http://arxiv.org/abs/1502.03622> (2015).
- [51] ———, *Metastability and higher-order computability*, Proceedings of LFCS18, Lecture Notes in Computer Science, Springer (2018).
- [52] Arthur Schöenflies, *Entwicklungen der Mengenlehre und ihrer Anwendungen, Part I, 2nd edition*, Leipzig-Berlin (Teubner), 1913.
- [53] Wilfried Sieg, *Hilbert's programs and beyond*, Oxford University Press, Oxford, 2013.
- [54] Waclaw Sierpiński, *Démonstration élémentaire du théorème sur la densité des ensembles*, Fund. Math. **4** (1923), 167–171.
- [55] Stephen G. Simpson, *Subsystems of second order arithmetic*, 2nd ed., Perspectives in Logic, CUP, 2009.
- [56] ———, *Friedman's research on subsystems of second-order arithmetic*, Harvey Friedman's research on the foundations of mathematics, Stud. Logic Found. Math., vol. 117, North-Holland, 1985, pp. 137–159.
- [57] Anne Sjerp Troelstra and Dirk van Dalen, *Constructivism in mathematics. Vol. I*, Studies in Logic and the Foundations of Mathematics, vol. 121, North-Holland, 1988.
- [58] Alan Turing, *On computable numbers, with an application to the Entscheidungs-problem*, Proceedings of the London Mathematical Society **42** (1936), 230–265.
- [59] Jean van Heijenoort, *From Frege to Gödel. A source book in mathematical logic, 1879–1931*, Harvard University Press, Cambridge, Mass., 1967.
- [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] Hermann Weyl, Edmund Landau, and Bernhard Riemann, *Das Kontinuum, und andere Monographien*, Chelsea, 1960.
- [62] W. H. Young and G. H. Young, *On The Reduction Of Sets Of Intervals*, Proc. Lond. Math. Soc. **14** (1915), 111–130.