LECTURE NOTES IN LOGIC

REVERSE MATHEMATICS 2001

EDITED BY

STEPHEN G. SIMPSON

CAMBRIDGE

ASL

LECTURE NOTES IN LOGIC 21

Reverse Mathematics 2001

Edited by

STEPHEN G. SIMPSON

Pennsylvania State University

ASSOCIATION FOR SYMBOLIC LOGIC

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

4843/24, 2nd Floor, Ansari Road, Daryaganj, Delhi – 110002, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781316755846
10.1017/9781316755846

Association for Symbolic Logic
Richard A. Shore, Publisher
Department of Mathematics, Cornell University, Ithaca, NY 14853
http://www.aslonline.org

A catalogue record for this publication is available from the British Library.

TABLE OF CONTENTS

POSSIBLE m-DIAGRAMS OF MODELS OF ARITHMETIC

ANDREW ARANA†

Abstract. In this paper we investigate the complexity of m-diagrams of models of various completions of first-order Peano Arithmetic (PA). We obtain characterizations that extend Solovay's results for open diagrams of models of completions of PA. We first characterize the m-diagrams of models of True Arithmetic by showing that the degrees of m-diagrams of nonstandard models $\mathcal{A}$ of TA are the same for all $m \geq 0$. Next, we obtain a more complicated characterization for arbitrary completions of PA. We then provide examples showing that some of the extra complication is needed. Lastly, we characterize sequences of Turing degrees that occur as $(\deg(T \cap \Sigma_n))_{n \in \omega}$, where T is a completion of PA.

§1. Introduction. We use $P(\omega)$ to denote the class of all subsets of ω. Let $\mathcal{L}_{\text{PA}}$ be the usual language of PA: relations $+$, $\cdot$, S, and $<$; and constants 0 and 1. We abbreviate True Arithmetic, the theory of the standard model of PA, by the initials TA. We use $S^n(0)$ to denote the numeral for n.

We continue with some preliminary definitions and results. A B_n formula is a boolean combination of Σ_n formulas. A *complete B_n type* is the set of all B_n formulas true of some tuple in some structure. The *open* diagram of a structure $\mathcal{A}$, denoted $D(\mathcal{A})$, is the collection of open sentences, with constants from $\mathcal{A}$, that are true in $\mathcal{A}$. Similarly, the m-diagram of $\mathcal{A}$, denoted $D_m(\mathcal{A})$, is the collection of B_m sentences, with constants from $\mathcal{A}$, that are true in $\mathcal{A}$.

Behind most of what we know about models and completions of PA is the notion of a Scott set:

Definition 1.1. A *Scott set* is a nonempty family of sets $\mathcal{S} \subseteq P(\omega)$ such that

1. if $X \in \mathcal{S}$ and $Y \leq_T X$, then $Y \in \mathcal{S}$,
2. if $X, Y \in \mathcal{S}$, then $X \oplus Y \in \mathcal{S}$,
3. if $T \subseteq 2^{<\omega}$ is an infinite tree in $\mathcal{S}$, then T has a path in $\mathcal{S}$. Equivalently, if A is a consistent set of sentences in $\mathcal{S}$, then some complete extension of A is in $\mathcal{S}$.

Received by the editor September 12, 2000.

†I am grateful to Julia Knight for helpful conversations and extraordinary patience concerning this paper.

Reverse Mathematics 2001
Edited by S. G. Simpson
Lecture Notes in Logic, 21

The family of arithmetical sets forms a Scott set. Scott sets are the ω-models of the axiom system WKL_0 as studied in reverse mathematics (and where the model is identified with the power set part of the structure, as in [14]). For a nonstandard model $\mathcal{A} \models$ PA, let $SS(\mathcal{A}) = \{d_a : a \in \mathcal{A}\}$, where

$$d_a = \{n \in \omega : \mathcal{A} \models p_n | a\}$$

where $(p_k)_{k \in \omega}$ is the sequence of primes.

Theorem 1.2 (Scott). *For a nonstandard model $\mathcal{A} \models$ PA, $SS(\mathcal{A})$ is a Scott set.*

We thus call $SS(\mathcal{A})$ the *Scott set* of the model $\mathcal{A}$.

The following well-known lemma is a sort of weak saturation property for bounded types in a Scott set:

Lemma 1.3. *Let $\mathcal{A}$ be a nonstandard model of* PA. *Let $\Gamma(\overline{u}, x)$ be a complete B_m type, with $\overline{a} \in \mathcal{A}$ a tuple that can be substituted for $\overline{u}$ in Γ. Then $\Gamma(\overline{a}, x)$ is realized by some $c \in \mathcal{A}$ if and only if $\Gamma(\overline{a}, x) \cup D_{m+1}(\mathcal{A})$ is consistent and $\Gamma(\overline{u}, x) \in SS(\mathcal{A})$.*

Scott was originally interested in Scott sets because they are closely tied to the notion of "representability". He wanted to characterize the families of sets representable with respect to completions of PA.

Definition 1.4. For a theory T in the language of PA, a set $X \subseteq \omega$ is *representable* by T if there is a formula φ such that for $n \in X$, $T \vdash \varphi(S^{(n)}(0))$, and for $n \notin X$, $T \vdash \neg\varphi(S^{(n)}(0))$.

We denote the collection of sets representable by a theory T by $\text{Rep}(T)$. Scott [12] showed the following fact relating Scott sets and $\text{Rep}(T)$:

Theorem 1.5 (Scott). *For a countable collection $\mathcal{S} \subseteq P(\omega)$, $\mathcal{S}$ is a Scott set if and only if there exists a completion T of* PA *such that* $\text{Rep}(T) = \mathcal{S}$.

Feferman [3] noted the following fact about nonstandard models of TA:

Theorem 1.6 (Feferman). *Let $\mathcal{A}$ be a nonstandard model of* TA. *Then $SS(\mathcal{A})$ contains the arithmetical sets.*

Feferman gave the result only for TA. However, for essentially the same reasons we also get the following result, for any model of PA:

Theorem 1.7. *Let $\mathcal{A}$ be a nonstandard model of* PA. *Then $SS(\mathcal{A})$ contains* $\text{Rep}(T)$. *Equivalently, $SS(\mathcal{A})$ contains $T_n = T \cap \Sigma_n$, for all n.*

Theorem 1.7 implies Theorem 1.6, because for $T =$ TA, $T_n \equiv_T \emptyset^{(n)}$ for all n. Theorem 1.7 suggests the following definition:

Definition 1.8. A Scott set $\mathcal{S}$ is *appropriate* for a theory T if $T_n \in \mathcal{S}$ for all n. Equivalently, $\mathcal{S}$ is appropriate for T if $\text{Rep}(T) \in \mathcal{S}$.

Using this definition, we can restate Theorem 1.7 as:

THEOREM 1.9. *Let $\mathcal{A}$ be a nonstandard model of* PA. *Then $SS(\mathcal{A})$ is appropriate for T.*

A notion we shall use in connection with Scott sets is that of an "enumeration".

DEFINITION 1.10. An *enumeration* of a set $\mathcal{S} \subseteq P(\omega)$ is a binary relation R such that $\mathcal{S} = \{R_n : n \in \omega\}$, where

$$R_n = \{k : (n, k) \in R\}.$$

An *R-index* for X is some $k \in \omega$ such that $R_k = X$.

DEFINITION 1.11. For a nonstandard model $\mathcal{A}$ of PA with universe ω,

$$R = \big\{(a, n) : \mathcal{A} \models p_n | a\big\}$$

is called the *canonical enumeration* of $SS(\mathcal{A})$.

We have the well-known fact:

PROPOSITION 1.12. *Let $\mathcal{A}$ be any nonstandard model of* PA *with universe ω and let R be the canonical enumeration of $SS(\mathcal{A})$. Then $R \leq_T D(\mathcal{A})$.*

This follows from the fact that the open diagram $D(\mathcal{A})$ witnesses true instances of the division algorithm. The following corollary follows from the fact that $D(\mathcal{A}) \leq_T D_m(\mathcal{A})$, for $m \geq 0$:

COROLLARY 1.13. *For $\mathcal{A}$ a nonstandard model of* PA *with universe ω, if R is the canonical enumeration of $SS(\mathcal{A})$, then $R \leq_T D_m(\mathcal{A})$, for $m \geq 0$.*

Solovay defined the notion of an "effective enumeration":

DEFINITION 1.14. For a countable Scott set $\mathcal{S}$, an *effective enumeration* is an enumeration R, with associated functions f, g, and h witnessing that $\mathcal{S}$ is a Scott set. These functions have the following properties:

1. if $\varphi_e^{R_i} = \chi_X$, then $f(i, e)$ is an R-index for X,
2. $g(i, j)$ is an R-index for $R_i \oplus R_j$,
3. if R_i is an infinite tree $T \subseteq 2^{<\omega}$, then $h(i)$ is an R-index for a set X such that χ_X is a path through T.

We say that an effective enumeration is *computable in a set X* if the enumeration and the three functions are all computable in X. Effective enumerations are available to us in light of the following result [7]:

THEOREM 1.15 (Marker). *Let $\mathcal{S}$ be a countable Scott set. If $\mathcal{S}$ has an enumeration computable in X, then it also has an effective enumeration computable in X.*

Solovay gave a characterization of the degrees (of open diagrams) of nonstandard models of TA in terms of effective enumerations [15]. Marker simplified Solovay's result by applying Theorem 1.15 [7]. The result is the following characterization:

Theorem 1.16 (Solovay/Marker). *The degrees of nonstandard models of TA are the degrees of enumerations of Scott sets containing the arithmetical sets.*

Solovay also characterized the degrees of (open diagrams of) nonstandard models of other completions of PA. The result is more difficult to state than the result for TA. To see why, let us highlight the difference between TA and arbitrary completions of PA. For a nonstandard model $\mathcal{A}$ of TA, $\mathcal{A}''$ yields the theory (and indices for the Σ_n fragments). For an arbitrary completion of PA this may not be so, as we will illustrate in Section 4.

Solovay found the general relationship between jumps of the model and indices for fragments of the theory. The result is the following characterization:

Theorem 1.17 (Solovay). *Suppose T is a completion of* PA. *The degrees of nonstandard models of T are the degrees of sets X such that*:

(a) *There is an enumeration $R \leq_T X$ of a Scott set $\mathcal{S}$ appropriate for T; and*
(b) *There are functions t_n for $n \geq 1$, $\Delta^0_n(X)$ uniformly in n, such that $\lim_{s\to\infty} t_n(s)$ is an index for T_n and for all s, $t_n(s)$ is an R-index for a subset of T_n.*

Solovay did not publish these results we are attributing to him. In [6], Julia Knight has given proofs of Theorems 1.16 and 1.17. Our proofs in Sections 2 and 3 follow those of [6], extending Solovay's results. In Section 2, we extend Solovay and Marker's characterization to include m-diagrams of nonstandard models of TA. In Section 3, we extend Solovay's characterization for arbitrary completions of PA to include m-diagrams. In Section 4, we will develop a class of theories $T(X)$ illustrating why the extra conditions in the more general characterization for arbitrary completions of PA given in Section 3 cannot simply be dropped. As part of doing this, we give a proof of Harrington's result that there exists a nonstandard model $\mathcal{A} \models$ PA such that $\mathcal{A} \leq_T 0'$ and $\mathrm{Th}(\mathcal{A})$ is not arithmetical [5]. Lastly, in Section 5, we examine the relationship between sequences of Turing degrees and completions of PA.

§2. True Arithmetic.

In this section we characterize the degrees of m-diagrams of nonstandard models of TA as the degrees of enumerations of Scott sets containing the arithmetical sets. We first show that for a nonstandard model of TA, we can find an enumeration below the m-diagram (in terms of Turing reducibility). We then show that for a suitable enumeration, we can find the m-diagram of a nonstandard model of TA below it. The second step requires more work. We use the fact that if R is an enumeration of a Scott set containing the arithmetical sets, then computably in R'' we can compute a sequence $(i_n)_{n\in\omega}$ of indices such that $R_{i_k} = \mathrm{TA} \cap \Sigma_k$ for each k. The fact holds because we can use R'' to list $\emptyset'$ and find its index in R; we may then use R'' to list $(\emptyset')'$, find its index in R, and so on. Using this fact, we can construct

Next, suppose $\mathcal{S}$ is a Scott set containing the arithmetical sets and that R is an enumeration of $\mathcal{S}$. We may use Marker's result again and take R to be an effective enumeration. To apply Theorem 2.3 and conclude the proof, we use a $\Delta_3^0(R)$ function $t(n)$ giving an R-index for $T_n = \mathrm{TA} \cap \Sigma_n$. Let $t(n)$ be the least R-index of $\mathrm{TA} \cap \Sigma_n$.

We show how to compute $t(n)$ using $\Delta_3^0(R)$. Note first that $\mathrm{TA} \cap \Sigma_n \leq_T \mathrm{TA} \cap \Sigma_{n+1}$ and $\mathrm{TA} \cap \Sigma_{n+1} \leq_T (\mathrm{TA} \cap \Sigma_n)'$ uniformly in n. Note also that the relation

$$J(i, j) = \{(i, j) : \forall x \left[x \in R_j \leftrightarrow x \in (R_i)'\right]\}$$

is $\Delta_3^0(R)$. Beginning with $t(r)$, an index for $\mathrm{TA} \cap \Sigma_r$, we use J to get an index for $(\mathrm{TA} \cap \Sigma_r)'$. Since $\mathrm{TA} \cap \Sigma_{r+1} \leq_T (\mathrm{TA} \cap \Sigma_r)'$, we use our effective enumeration to get an index for $\mathrm{TA} \cap \Sigma_{r+1}$. This index is $t(r+1)$.

We have thus shown $t(n)$ to be $\Delta_3^0(R)$. We may now apply Theorem 2.3 to get a nonstandard model $\mathcal{A}$ of TA such that $SS(\mathcal{A}) = \mathcal{S}$ and $D_m(\mathcal{A}) \leq_T R$. ⊣

As a corollary to the previous result, we have the following:

COROLLARY 2.7. *The degrees of m-diagrams of nonstandard models $\mathcal{A}$ of* TA *are the same for all $m \geq 0$.*

§3. Other completions of PA. In this section we give a characterization of the m-degrees of nonstandard models of other completions of PA. This new characterization (Theorem 3.4) will be like the characterization for TA (Theorem 2.6) in that it involves enumerations of an appropriate Scott set. It differs from the earlier characterization in that it additionally involves a sequence of approximating functions.

To prove this characterization, we need to use the sequence of oracles $(\Delta_i^0(X))_{i \in \omega}$ to prove a more general version of Theorem 2.1. To prove this result, Theorem 3.1, we use a infinitely nested priority construction. The result for $m = 1$ is Theorem 2.3 in [6]. Again, the proof for arbitrary m is essentially the same, so we omit details and give only a sketch.

As with the TA case, we can break the characterization into two parts. The model-construction part, Theorem 3.2, can itself again be separated into two separate priority constructions. The first priority construction for TA, Theorem 2.1, used $\Delta_2^0(X)$ to approximate a $\Delta_3^0(X)$ function. In the case of arbitrary completions of PA, we need to approximate not a single $\Delta_3^0(X)$ function, but rather a sequence of functions t_{m+n}, $\Delta_n^0(X)$ uniformly in n, approximating $T \cap \Sigma_n$ for each n relative to X. We thus need to prove a more general version of Theorem 2.1. Here we use an infinitely nested priority construction.

Infinitely nested priority constructions are difficult to do in general. However, there is a metatheorem giving conditions under which some may be done. Solovay's theorem and our generalization follow from the metatheorem 4.1 in [6].

As with TA, our plan is to build a nonstandard model $\mathcal{B}$ such that $D_m(\mathcal{B}) \leq_T X''$ and such that the set

$$Q = \{(i, \overline{a}) : R_i \text{ is the complete } B_{m+1} \text{ type of } \overline{a}\}$$

is $\Sigma^0_2(X)$. The metatheorem shows that under certain conditions such a construction can be effected.

The result of this construction is the following:

THEOREM 3.1. *Let T be a completion of* PA *and let $m \geq 0$. Suppose $R \leq_T X$ is an enumeration of a Scott set $\mathcal{S}$, with functions t_{m+n} for $n \geq 2$, $\Delta^0_n(X)$ uniformly in n, such that $\lim_{s\to\infty} t_{m+n}(s)$ is an index for T_{m+n} and for all s, $t_{m+n}(s)$ is an index for a subset of T_{m+n}. Then T has a model $\mathcal{A}$ such that $SS(\mathcal{A}) = \mathcal{S}$ and $Q = \{(i, \overline{a}) : R_i$ is the complete B_{m+1} type of $\overline{a}\}$ is $\Sigma^0_2(X)$.*

We may now reuse Theorem 2.2, using $\Delta^0_1(X)$ to approximate $\mathcal{B}$, building an isomorphic copy $\mathcal{A}$ such that $D_m(\mathcal{A}) \leq_T X$. These constructions can then be combined into one result:

THEOREM 3.2. *Let T be a complete theory and let $m \geq 0$. Suppose $R \leq_T X$ is an enumeration of a Scott set $\mathcal{S}$, with functions t_{m+n} for $n \geq 2$, $\Delta^0_n(X)$ uniformly in n, such that $\lim_{s\to\infty} t_{m+n}(s)$ is an index for T_{m+n} and for all s, $t_{m+n}(s)$ is an index for a subset of T_{m+n}. Then T has a model $\mathcal{A}$ with $SS(\mathcal{A}) = \mathcal{S}$ such that $D_m(\mathcal{A}) \leq_T X$.*

To show the enumeration half of the main theorem, we need a modified version of Solovay's Approximation Lemma for m-diagrams. The original version for $m = 1$ appears in [6] along with a proof. We omit the details here, as the proof for arbitrary m is essentially the same.

LEMMA 3.3. *Let $\mathcal{A}$ be a nonstandard model of* PA *with universe ω, and let R be the canonical enumeration of $SS(\mathcal{A})$. Then for any $m \geq 0$, there are functions t_{m+n}, $\Delta^0_n(D_m(\mathcal{A}))$ uniformly in n, such that $\lim_{s\to\infty} t_{m+n}(s)$ is an R-index for $T_{m+n}(\mathcal{A})$. Furthermore, for $r < s$, $R_{t_n(r)} \subseteq R_{t_n(s)}$.*

We can now give the main result giving the characterization for an arbitrary completion of PA:

THEOREM 3.4. *Suppose T is a completion of* PA. *For any $m \geq 0$, the degrees of m-diagrams of nonstandard models of T are the degrees of sets X such that:*

(a) *There is an enumeration $R \leq_T X$ of a Scott set $\mathcal{S}$ appropriate for T; and*
(b) *There are functions t_{m+n} for $n \geq 1$, $\Delta^0_n(X)$ uniformly in n, such that $\lim_{s\to\infty} t_{m+n}(s)$ is an index for T_{m+n} and for all s, $t_{m+n}(s)$ is an R-index for a subset of T_{m+n}.*

PROOF. Suppose first that $R \leq_T X$ is an enumeration $\mathcal{S}$ satisfying condition (b) above. Using Theorem 3.2, we get a model $\mathcal{A} \models T$ with $SS(\mathcal{A}) = \mathcal{S}$ such that $D_m(\mathcal{B}) \leq_T R$. Next, suppose we start with $\mathcal{A} \models T$ with $SS(\mathcal{A}) = \mathcal{S}$ such that $D_m(\mathcal{B}) \leq_T X$. Using the canonical enumeration R of $SS(\mathcal{A})$, we get that $R \leq_T D_m(\mathcal{A})$. Then by Lemma 3.3, functions satisfying (b) exist as needed. ⊣

§4. Examples. In this section, we present examples illustrating aspects of Solovay's results. First, we give a theory T with enumeration R of $\text{Rep}(T)$ such that there is no model of T computable in R. Next, we present Harrington's result that there is a model $\mathcal{A}$ of PA that is computable in $0'$, but $\text{Th}(\mathcal{A})$ is not arithmetical. Hence, $\text{Th}(\mathcal{A}) \not\leq_T \mathcal{A}^{(n)}$ for any n. Thus, Solovay's results in general require an infinite sequence of approximating functions. In this sense especially, arbitrary completions of PA differ from TA.

We provide a general procedure for constructing the theories we use in these examples in Theorem 4.4. The construction uses the Gödel-Rosser Incompleteness Theorem, as well as Scott's modification of this theorem. We will review the Gödel-Rosser and Scott results before giving our results.

Independence was first explored by Gödel in his landmark 1931 paper [4]. Rosser tightened the result by modifying the sentence shown to be independent [11]. We state a variant of the Gödel-Rosser Theorem that we will make use of later:

Lemma 4.1 (Gödel-Rosser). *There is a computable sequence of sentences* $(\varphi_n)_{n\in\omega}$ *such that* φ_n *is* Π_{n+1} *and for any set* Γ *of* B_n *sentences consistent with* PA, φ_n *is independent over* $\text{PA} \cup \Gamma$.

Note that we may also extend the axioms of PA by any computable set and preserve the result.

We continue with Scott's results. In arriving at his results regarding Scott sets, Scott investigated the notion of independence for formulas.

Definition 4.2. For a set of sentences Γ and a formula $\varphi(x)$, $\varphi(x)$ is *independent* over Γ if for all $X \subseteq \omega$, the set

$$\Gamma \cup \left\{\varphi\left(S^{(n)}(0)\right) : n \in X\right\} \cup \left\{\neg\varphi\left(S^{(n)}(0)\right) : n \notin X\right\}$$

is consistent.

By varying the Gödel-Rosser independent sentence, Scott was able to show the following result [12]:

Lemma 4.3 (Scott). *There is a computable sequence of formulas* $(\varphi_n(x))_{n\in\omega}$ *such that* φ_n *is* Π_{n+2} *and if* Γ *is a set of* B_n *sentences such that* $\text{PA} \cup \Gamma$ *is consistent, then* φ_n *is independent over* $\text{PA} \cup \Gamma$.

Let's consider briefly the construction of these independent formulas. Fix n. We sketch the construction of the formula $\varphi_n(x)$ in two steps. The first step is to define a sequence of Π_{n+1} sentences $(\psi_\sigma)_{\sigma\in 2^{<\omega}}$, which we think of as being on a binary-branching tree τ. We describe the first few levels of τ. At level 0 of τ, let the root be a variant of the Gödel-Rosser sentence that says "for any proof of me from PA and true B_n sentences, there is a smaller proof of my negation from the same axioms"; call this sentence $\psi_{\langle\emptyset\rangle}$. The root $\psi_{\langle\emptyset\rangle}$ branches left to a sentence $\psi_{\langle 0\rangle}$ that says, "for any proof of me from PA, true B_n sentences, and $\psi_{\langle\emptyset\rangle}$, there is a smaller proof of my negation from the same axioms". Similarly, $\psi_{\langle\emptyset\rangle}$ branches right to $\psi_{\langle 1\rangle}$, which says, "for any proof of me from PA, true B_n sentences, and $\neg\psi_{\langle\emptyset\rangle}$, there is a smaller proof of my negation from the same axioms". Both $\psi_{\langle 0\rangle}$ and $\psi_{\langle 1\rangle}$ are at level 1 of τ. We may continue and define the level 2 sentences of τ similarly: $\psi_{\langle 0\rangle}$ branches to the left to a sentence $\psi_{\langle 00\rangle}$ that says "for any proof of me from PA, true B_n sentences, $\psi_{\langle\emptyset\rangle}$, and $\psi_{\langle 0\rangle}$, there is a smaller proof of my negation from the same axioms", while $\psi_{\langle 0\rangle}$ branches to the right to a sentence $\psi_{\langle 01\rangle}$ that says "for any proof of me from PA, true B_n sentences, $\psi_{\langle\emptyset\rangle}$, and $\neg\psi_{\langle 0\rangle}$, there is a smaller proof of my negation from the same axioms". Accordingly, $\psi_{\langle 1\rangle}$ branches to sentences $\psi_{\langle 10\rangle}$ and $\psi_{\langle 11\rangle}$. For each $\sigma \in 2^{<\omega}$, the sentence ψ_σ is defined as above, using σ to determine which axioms ψ_σ mentions. Each sentence ψ_σ is independent over PA, Γ, and the axioms $\pm\psi_\varsigma$ that ψ_σ mentions.

Using this sequence $(\psi_\sigma)_{\sigma\in 2^{<\omega}}$ of Π_{n+1} sentences, we specify another sequence of sentences $(\mu_n)_{n\in\omega}$. Each sentence μ_i expresses the disjunction of all paths of length $i+1$ through τ that branch to the left at level i. We illustrate this by giving the first three sentences of this sequence. First, let

$$\mu_0 = \psi_{\langle\emptyset\rangle}.$$

Next, let

$$\mu_1 = \big(\psi_{\langle\emptyset\rangle} \wedge \psi_{\langle 0\rangle}\big) \vee \big(\neg\psi_{\langle\emptyset\rangle} \wedge \psi_{\langle 1\rangle}\big).$$

Continuing, let

$$\begin{aligned}\mu_2 = {} & \big(\psi_{\langle\emptyset\rangle} \wedge \psi_{\langle 0\rangle} \wedge \psi_{\langle 00\rangle}\big) \vee \big(\psi_{\langle\emptyset\rangle} \wedge \neg\psi_{\langle 0\rangle} \wedge \psi_{\langle 01\rangle}\big) \vee \\ & \big(\neg\psi_{\langle\emptyset\rangle} \wedge \psi_{\langle 1\rangle} \wedge \psi_{\langle 10\rangle}\big) \vee \big(\neg\psi_{\langle\emptyset\rangle} \wedge \neg\psi_{\langle 1\rangle} \wedge \psi_{\langle 11\rangle}\big).\end{aligned}$$

Continue this way for all levels i. Since these sentences μ_n are boolean combinations of Π_{n+1} sentences, each μ_n may be taken to be B_{n+1}.

We are now finally ready to describe the formula $\varphi_n(x)$ described in the lemma. Let $\varphi_n(x) = \mathrm{Sat}_{B_{n+1}}(\mu_x)$. We may take $\mathrm{Sat}_{B_{n+1}}(x)$ to be both Π_{n+2} and Σ_{n+2}.

We will use Lemmas 4.1 and 4.3 for our examples, by way of the following construction. We remark that Marker proved essentially the same result in his Ph.D. thesis [9], using essentially the same proof. The result appears there as Theorem 1.27.

THEOREM 4.4. *Let R be an enumeration of a Scott set $\mathcal{S}$. For any set X, there is a completion $T(X,R)$ of* PA *with* $\mathrm{Rep}(T(X,R)) = \mathcal{S}$ *and* $T(X,R) \cap B_{3n} \leq_T (X \cap n) \oplus R$, *uniformly in n.*

PROOF. We may suppose R is an effective enumeration, by Marker's Theorem 1.15. We construct the appropriate theory $T(X)$. We start with a computable sequence $(\varphi_n(x))_{n\in\omega}$ of independent formulas as in Lemma 4.3, where $\varphi_n(x)$ is Π_{n+2}. We also start with a computable sequence $(\varphi_n^*)_{n\in\omega}$ of independent sentences as in Lemma 4.1, where φ_n^* is Π_{n+1}. Let T be any completion of PA. We build $T(X,R)$ using the following list of requirements:

Code$_0$: Take the Π_1 sentence φ_0^* from the sequence given by Lemma 4.1, where φ_0^* is independent over PA.

If $0 \in X$, let $T_1^* =$ a completion of PA $\cup \{\varphi_0^*\}$ in $\mathcal{S}$.

If $0 \notin X$, let $T_1^* =$ a completion of PA $\cup \{\neg\varphi_0^*\}$ in $\mathcal{S}$.

We may do this because φ_0^* and $\neg\varphi_0^*$ are both consistent with PA $\cup (T \cap B_0)$. In either case we can effectively find the index i_1^* of the completion.

Let $T_1 = T_1^* \cap B_1$. We can find its index i_1 effectively as well. Informally, T_1 "codes" whether or not $0 \in X$.

Code$_1$: Take the Π_3 formula $\varphi_1(x)$ from the sequence given by Scott's Lemma 4.3, where $\varphi_1(x)$ is independent over PA $\cup$ T_1. For $k \in R_0$, put $\varphi_1(S^{(k)}(0))$ into T_3^*. For $k \notin R_0$, put $\neg\varphi_1(S^{(k)}(0))$ into T_3^*.

Next, we find the index for a completion of PA $\cup$ $T_1 \cup \{\varphi_1(S^{(k)}(0)) : k \in R_0\} \cup \{\neg\varphi_1(S^{(k)}(0)) : k \notin R_0\}$. Then let T_3 be the B_3 part of this completion, again finding its index i_3. Informally, T_3 codes that R_0 is in Rep(T).

Code$_{2n}$: Take the Π_{3n+1} sentence φ_{3n}^*, where φ_{3n}^* is independent over PA $\cup$ T_{3n}.

If $n \in X$, let $T_{3n+1}^* =$ a completion of PA $\cup (T_{3n} \cap B_{3n}) \cup \{\varphi_{3n}^*\}$ in $\mathcal{S}$.

If $n \notin X$, let $T_{3n+1}^* =$ a completion of PA $\cup (T_{3n} \cap B_{3n}) \cup \{\neg\varphi_n^*\}$ in $\mathcal{S}$.

Once again, we can effectively find the index i_{3n+1}^* of T_{3n+1}^*. Let $T_{3n+1} = T_{3n+1}^* \cap B_{3n+1}$. We can find its index i_{3n+1} effectively as well.

Code$_{2n+1}$: Take the Π_{3n+3} formula $\varphi_{3n+1}(x)$ of our sequence, where $\varphi_{3n+1}(x)$ is independent over PA $\cup$ T_{3n+1}. For $k \in R_n$, put $\varphi_{3n+1}(S^{(k)}(0))$ into T_{3n+3}^*. For $k \notin R_n$, put $\neg\varphi_{3n+1}(S^{(k)}(0))$ into T_{3n+3}^*.

Next, we find an index for a completion of

$$\text{PA} \cup T_{3n+1} \cup \left\{\varphi_{3n+1}\left(S^{(k)}(0)\right) : k \in R_n\right\} \cup \left\{\neg\varphi_{3n+1}\left(S^{(k)}(0)\right) : k \notin R_n\right\}.$$

Then let T_{3n+3} be the B_{3n+3} part of this completion, finding its index i_{3n+3}.

This ends our inductive definition of $T(X,R)$. By our construction, it is clear that Rep$(T(X,R)) = S$. ⊣

Note that our construction also gives that $X \leq_T T(X,R)$. To determine if $n \in X$, we may ask $T(X,R)$ which of $\pm\varphi_{3n}^* \in T(X,R)$. If $\varphi_{3n}^* \in T(X,R)$, then $n \in X$; if $\neg\varphi_{3n}^* \in T(X,R)$, then $n \notin X$.

We can use this construction to build the following theory, demonstrating that the extra conditions requiring approximating functions for the fragments of the theory in Theorems 1.17 and 3.4 cannot be dropped:

Corollary 4.5. *For any enumeration R of a Scott set $\mathcal{S}$, there is a completion T of* PA *such that* Rep$(T) = \mathcal{S}$ *and there is no model $\mathcal{A} \models T$ such that $\mathcal{A} \leq_T R$.*

Proof. Let X be a set such that $X \not\leq_T R^{(\omega)}$. Let T be a completion given by the construction of Theorem 4.4. We show that if $\mathcal{A} \models T$, then $\mathcal{A} \not\leq_T R$. If denoted τ_n, as an infinite binary-branching tree as follows. A node $\sigma \in 2^{<\omega}$ is in τ_n iff there is no proof of a contradiction of length less than $\mathrm{len}(\sigma)$ from the set

$$\mathrm{PA} \cup \widetilde{T_n} \cup \{\varphi_k : \sigma(k) = 1\} \cup \{\neg\varphi_k : \sigma(k) = 0\}.$$

Each path through τ_n corresponds to a completion of $\mathrm{PA} + \widetilde{T_n}$, since paths decide every sentence from $(\varphi_k)_{k\in\omega}$ consistently.

Fix $n \geq 0$. Note that $\tau_n \leq_T T_n$. By Lemmas 5.3 and 5.4, there is a path T^* through τ_n computable in T_{n+1}. Note that we have that $T_n = T_n^*$ and $T_n^* \leq_T T^*$. Then, since $\mathrm{Rep}(T^*)$ is a Scott set, we have that $T_n^* \in \mathrm{Rep}(T^*)$. Hence $T_n \in \mathrm{Rep}(T^*)$. ⊣

Next, we prove Lemmas 5.3 and 5.4. First, we prove Lemma 5.3:

Proof. Fix n. Since $\tau \leq_T T_n$, there is some e such that $\chi_\tau(x) = \varphi_e^{T_n}(x)$. It is well-known that there are Π_1 and Σ_1 formulas $\psi_\Pi(e, x, y, z, \sigma)$ and $\psi_\Sigma(e, x, y, z, \sigma)$ representing that y is a computation of φ_e on input x using oracle σ with output y. To represent the oracle T_n, we make use of the formula $\mathrm{Sat}_n(x)$, defining truth for Σ_n sentences. It is well-known that $\mathrm{Sat}_n(x)$ is Σ_n. Using $\mathrm{Sat}_n(x)$, we give a Σ_{n+1} formula for representing $\varphi_e^{T_n}(x)$:

$$\begin{aligned}\delta_\Sigma(x) = \exists y\, \exists\sigma\, \big[\psi_\Sigma(e, x, y, 1, \sigma) \\ \wedge\forall t < \mathrm{len}(\sigma)\, \big[(\sigma(t) = 1 \rightarrow \mathrm{Sat}_n(t)) \\ \wedge(\sigma(t) = 0 \rightarrow \neg\mathrm{Sat}_n(t))\big]\big].\end{aligned}$$

We also give a Π_{n+1} representing formula:

$$\begin{aligned}\delta_\Pi = \forall y\, \forall\sigma\, \big[\psi_\Sigma(e, x, y, 0, \sigma) \\ \rightarrow \exists t < \mathrm{len}(\sigma)\, \big[(\sigma(t) = 1 \wedge \neg\mathrm{Sat}_n(t)) \\ \vee(\sigma(t) = 0 \wedge \mathrm{Sat}_n(t))\big]\big].\end{aligned}$$ ⊣

We now prove Lemma 5.4:

Proof. There are two cases to consider. In Case 1, we assume that T proves that τ has an infinite path. In Case 2, we assume that T proves that τ does not have an infinite path. In both cases, we show how to find ζ, a path through τ.

Case 1: T proves that τ has an infinite path.

We first present a Π_{n+1} formula infinite-left(τ) that holds iff a node $\sigma \in \tau$ has an infinite extension in τ to its left:

$$\begin{aligned}\text{infinite-left}(\sigma) := \forall s > \mathrm{len}(\sigma)\, \exists\gamma \\ \leq\, p(s+1)\big[(\mathrm{len}(\gamma) = s+1) \\ \wedge((\sigma \wedge 0) \subseteq \gamma) \wedge \delta_\Pi(\gamma)\big].\end{aligned}$$

Suppose we have determined an initial segment σ_i of our path ζ through τ, where σ_i has length i. Here is how we decide whether to branch to the left or right at the ith level in our path. We update our path to $\sigma_i \wedge 0$ if infinite-left$(\sigma_i) \in T_{n+1}$. We update our path to $\sigma_i \wedge 1$ if $\neg$ infinite-left$(\sigma_i) \in T_{n+1}$. We do this for every $i \geq 0$. Let

$$\zeta = \bigcup_{i \in \omega} \sigma_i.$$

Since we use T_{n+1} as an oracle, we get that $\zeta \leq_T T_{n+1}$.

CASE 2: T does not prove that τ has an infinite path.

In this case, T proves that there is some level past which no node in the tree can be consistently extended. We extend an initial segment to this maximum level, and take this initial segment to be our ζ. Since we will use T_{n+1} as an oracle, we will have $\zeta \leq_T T_{n+1}$.

Suppose we have determined an initial segment σ_i of our path ζ through τ, where σ_i has length i. Here is how we decide whether to branch to the left or right at the ith level in our path. Since we are in Case 2, T witnesses that τ is finite. Thus T proves that there is some first level s_0 to the left of σ_i and some first level s_1 to the right of σ_i, beyond which no path can be consistently extended. We extend to $\sigma_i \wedge 0$ if $s_0 > s_1$, while we extend to $\sigma_i \wedge 1$ if $s_1 \geq s_0$.

To decide whether $s_0 > s_1$ or $s_1 \geq s_0$, we use two Σ_{n+1} formulas, $\psi_0(\pi)$ and $\psi_1(\pi)$. The formula $\psi_0(\pi)$ holds if there is a level s such that there is some node extending the node π to the left that is contained in τ; while at the same time, there is no node at level s extending π to the right that is contained in τ. The formula ψ_1 is similar but considers extensions to the right. Here are the formulas:

$$\begin{aligned}\psi_0(\pi) := \exists s\, \exists \sigma \leq p(s+1)\, &(\mathrm{len}(\sigma) = s+1) \\ &\wedge ((\pi \wedge 0) \subseteq \sigma) \wedge \delta_\Sigma(\sigma) \\ &\wedge \forall \lambda \leq p(s+1)\, \big[(\mathrm{len}(\lambda) = s+1) \\ &\qquad\qquad \wedge ((\pi \wedge 1) \subseteq \lambda) \rightarrow \neg\delta_\Pi(\lambda)\big]\end{aligned}$$

and

$$\begin{aligned}\psi_1(\pi) := \exists s\, \exists \sigma \leq p(s+1)\, \big[\big[&(\mathrm{len}(\sigma) = s+1) \\ &\wedge ((\pi \wedge 1) \subseteq \sigma) \wedge \delta_\Sigma(\sigma)\big] \\ &\wedge \forall \lambda \leq p(s+1)\, \big[(\mathrm{len}(\lambda) = s+1) \\ &\qquad \wedge ((\pi \wedge 0) \subseteq \lambda) \\ &\qquad \rightarrow \neg\delta_\Pi(\lambda)\big]\big].\end{aligned}$$

If $\psi_0(\tau_t) \in T_{n+1}$, then $s_0 > s_1$. If $\psi_1(\tau_t) \in T_{n+1}$, then $s_1 \geq s_0$. If neither is in T_{n+1} for some level i^*, then we have reached the maximum extendible level of τ, according to T_{n+1}. Let

$$T^* = \bigcup_{i < i^*} \sigma_i.$$

Since we use T_{n+1} as an oracle, we get that $T^* \leq_T T_{n+1}$. ⊣

This completes the proof of the (1) $\Rightarrow$ (2) direction of Theorem 5.1. Next, we give the (2) $\Rightarrow$ (1) direction, with proof:

THEOREM 5.5. *Suppose* $(\boldsymbol{d_n})_{n\in\omega}$ *is a sequence of Turing degrees such that*

$$\boldsymbol{0} = \boldsymbol{d_0} \ll \boldsymbol{d_1} \ll \boldsymbol{d_2} \ll \cdots .$$

Then there exists a completion T *of* PA *such that for all* n, $\boldsymbol{d_n} = \deg(T_n)$.

PROOF. We build the completion T inductively by determining each of its fragments T_i. Let $T_0 = \text{PA} \cap \Sigma_0$. For our inductive step, suppose we have specified T_{n-1}. We build T_n so that $T_n \equiv_T D_n$, for D_n a fixed representative from $\boldsymbol{d_n}$. After we show how to construct T_n, we will show that $T_n \equiv_T D_n$, by how we have constructed T_n.

By assumption, there is a completion T^* such that $T^* \leq_T D_n$ and $T_{n-1} \in \text{Rep}(T^*)$. Let φ_k be a computable list of the Σ_n sentences of $\mathcal{L}_{\text{PA}}$. We break our construction into attempts to meet the following requirements, for $k \geq 0$:

R_{2k}: Put one of φ_k or $\neg\varphi_k$ into T_n.

R_{2k+1}: Code whether $k \in D_n$ into T_n.

To meet these requirements, we define sets A_i such that

$$\bigcup_{i\in\omega} A_i = T_n.$$

First, let $A_0 = \text{PA} \cup T_{n-1}$. Suppose we have already defined A_j. There are two cases to consider:

CASE 1: j is even.

Then $j = 2k$, for some $k \geq 0$. We define A_{2k+1}, in an attempt to meet requirement R_{2k}. To decide whether to add φ_k or $\neg\varphi_k$ to T_n, we use the following notion. We say that $A_{2k} \cup \{\varphi_k\}$ is *more inconsistent* than $A_{2k} \cup \{\neg\varphi_k\}$, according to T^*, iff T^* proves that there is a smaller proof of an inconsistency from $A_{2k} \cup \{\varphi_k\}$ than there is from $A_{2k} \cup \{\neg\varphi_k\}$. Let γ_k be the sentence in $\mathcal{L}_{\text{PA}}$ expressing that $A_{2k} \cup \{\varphi_k\}$ is more inconsistent than $A_{2k} \cup \{\neg\varphi_k\}$.

CLAIM 1: If $\gamma_k \in T^*$, then $A_{2k} \cup \{\neg\varphi_k\}$ is consistent.

If we are in this case, then we put $\neg\varphi_k$ into T_n. Let $A_{2k+1} := A_{2k} \cup \{\neg\varphi_k\}$.

CLAIM 2: If $\gamma_k \notin T^*$, then $A_{2k} \cup \{\varphi_k\}$ is consistent.

If we are in this case, then we put φ_k into T_n. Let $A_{2k+1} := A_{2k} \cup \{\varphi_k\}$.

To finish describing how to meet the even requirements, we need to prove Claims 1 and 2. We leave those proofs until the end.

CASE 2: j is odd.

Then $j = 2k + 1$, for some $k \geq 0$. We build A_{2k+2}, in an attempt to meet requirement R_{2k+1} Use Lemma 4.1, the variant of the Gödel-Rosser Theorem, to get a Π_n sentence ψ_k, independent over A_{2k+1}.

If $k \in D_n$, put $\neg\psi_k$ into T_n. Let $A_{2k+2} = A_{2k+1} \cup \{\neg\psi_k\}$.

If $k \notin D_n$, put ψ_k into T_n. Let $A_{2k+2} = A_{2k+1} \cup \{\psi_k\}$.

This ends our description of the construction. No injury ever threatens these requirements, so in the limit they will all be met. Let $\bigcup_{i\in\omega} A_i = T_n$.

To show that $T_n \equiv_T D_n$, we must show both that $D_n \leq_T T_n$ and $T_n \leq_T D_n$. First, we show that $D_n \leq_T T_n$. To do this, we need to decode if $k \in D_n$, computably in T_n, by following the construction through requirement R_{2k+1}. We reuse the computable list $(\varphi_k)_{k\in\omega}$ of Σ_n sentences of $\mathcal{L}_{\mathrm{PA}}$. To begin decoding, ask T_n if $\varphi_0 \in T_n$. Using the answer, update the set A_1. Using A_1, we may use Lemma 4.1 to compute the Π_n sentence ψ_0 that is independent over A_1. Using T_n, we check whether $\pm\psi_0 \in T_n$. If $\psi_0 \in T_n$, then we know that $0 \notin D_n$, by construction. If $\neg\psi_0 \in T_n$, then we know $0 \in D_n$. In either case, we have decoded whether or not $0 \in D_n$. Use this answer to update A_2.

At step $2k$, ask T_n if $\pm\varphi_k$ is in T_n, as described above for step 0. Update A_{2k+1}. Do step $2k+1$, deciding if $\pm\psi_k$ in T_n as above for step 1, and hence decoding whether $k \in D_n$.

Next, we show that $T_n \leq_T D_n$. For a B_n sentence α, we want to determine whether $\alpha \in T_n$. By assumption, we have that there is a completion T^* such that $T^* \leq_T D_n$. We may follow through the steps of the construction given above, computably in D_n. At each even step $2k$, building A_{2k+1}, we check if $\alpha = \pm\varphi_k$. If it is, by following through the steps in Case 1, we determine whether or not we put $\alpha \in A_{2k+1}$. If it is not, we follow through the steps of Cases 1 and 2, reaching the next even step. Since our computable list $(\varphi_i)_{i\in\omega}$ contains every Σ_k sentence in $\mathcal{L}_{\mathrm{PA}}$, we will eventually reach an even step $2k$ where $\alpha = \pm\varphi_k$.

Finally, we give the proofs of Claims 1 and 2.

Proof of Claim 1: Suppose $\gamma_k \in T^*$. Let p witness γ_k in T^*. Then p is a proof of $\perp$ from $A_{2k} \cup \{\varphi_k\}$ in T^*, and for all $q < p$, q is not a proof of $\perp$ from $A_{2k} \cup \{\neg\varphi_k\}$ in T^*. If p is standard, then there is no standard proof of $\perp$ from $A_{2k} \cup \{\neg\varphi_k\}$, so $A_{2k} \cup \{\neg\varphi_k\}$ is consistent. If p is nonstandard, then since there is no smaller proof of $\perp$ from $A_{2k} \cup \{\neg\varphi_k\}$, in particular there can be no *standard* proof of $\perp$ from $A_{2k} \cup \{\neg\varphi_k\}$. Again, $A_{2k} \cup \{\neg\varphi_k\}$ is consistent.

Proof of Claim 2: Suppose $\gamma_k \notin T^*$. If there is no proof of $\perp$ from $A_{2k} \cup \{\varphi_k\}$, then we are finished. Suppose p is a proof of $\perp$ from $A_{2k} \cup \{\varphi_k\}$. If p is standard, then there is a proof $q < p$ of $\perp$ from $A_{2k} \cup \{\neg\varphi_k\}$. We then have that $A_{2k} \vdash \neg(\varphi_k \vee \neg\varphi_k)$, or equivalently, $A_{2k} \vdash \varphi_k \wedge \neg\varphi_k$. This contradicts the fact that we have built A_{2k} to be consistent. Thus p cannot be standard, so $A_{2k} \cup \{\varphi_k\}$ is consistent. ⊣

REFERENCES

[1] Andrew Arana, *Solovay's theorem cannot be simplified*, ***Annals of Pure and Applied Logic***, vol. 112 (2001), pp. 27–41.

[2] Jon Barwise (editor), ***Handbook of mathematical logic***, Studies in Logic and the Foundations of Mathematics, vol. 90, North-Holland Publishing Co., Amsterdam, 1977.

[3] Solomon Feferman, *Arithmetically definable models of formalizable arithmetic*, ***Notices of the American Mathematical Society***, vol. 5 (1958), p. 679.

[4] Kurt Gödel, *Über formal unentscheidhare Sätze der* Principia Mathematica *und verwandter Systeme I*, ***Collected works*** (Solomon Feferman et al., editors), vol. 1, Oxford University Press, New York and Oxford, 1986, pp. 145–195.

[5] Leo Harrington, *Building non-standard models of Peano Arithmetic*, Unpublished notes, 1979.

[6] Julia F. Knight, *True approximations and models of arithmetic*, ***Models and computability*** (B. Cooper and J. Truss, editors), Cambridge University Press, New York, 1999, pp. 255–278.

[7] Angus Macintyre and David Marker, *Degrees of recursively saturated models*, ***Transactions of the American Mathematical Society***, vol. 282 (1984), pp. 539–554.

[8] David Marker, *Degrees of models of true arithmetic*, ***Proceedings of the Herbrand Symposium*** (J. Stern, editor), North-Holland, 1982, pp. 233–242.

[9] ———, ***Degrees of models of arithmetic***, Ph.D. thesis, Yale University, 1983.

[10] Yuri Matijasevich, *Enumerable sets are Diophantine*, ***Doklady Akademii Nauka SSSR***, vol. 191 (1970), pp. 272–282, Russian.

[11] J. B. Rosser, *Extensions of some theorems of Gödel and Church*, ***The Journal of Symbolic Logic***, vol. 1 (1936), pp. 87–91.

[12] Dana Scott, *Algebras of sets binumerable in complete extensions of arithmetic*, ***Recursive function theory*** (J. Dekker, editor), Proceedings of Symposia in Pure Mathematics, vol. V, American Mathematical Society, 1962, pp. 117–122.

[13] Stephen G. Simpson, *Degrees of unsolvability: A survey of results*, [2], pp. 631–652.

[14] ———, ***Subsystems of second order arithmetic***, Perspectives in Mathematical Logic, Springer, 1999.

[15] Robert Solovay, *Degrees of models of true arithmetic*, Circulated preprint, 1982.

DEPARTMENT OF PHILOSOPHY
201 DICKENS HALL
KANSAS STATE UNIVERSITY
MANHATTAN, KS 66503-0803, USA
E-mail: aarana@ksu.edu
URL: http://www.stanford.edu/~aarana

The set of primitive recursive functions is the smallest set of functions of various arities from the natural numbers to the natural numbers, containing the constant zero, projections, and the successor function, and closed under composition and primitive recursion. The language of *PRA* has a symbol for each primitive recursive function, and the axioms of *PRA* consist of quantifier-free defining equations for these functions and a schema of induction for quantifier-free formulae.

Identifying relations with their characteristic functions, one can use primitive recursion to define the relation $x < y$; or, equivalently, one can add a relation symbol to the language of *PRA* with appropriate defining equations. The schema of induction is equivalent to

$$\forall x\ (\varphi(0) \wedge \forall y < x\ (\varphi(y) \rightarrow \varphi(y+1)) \rightarrow \varphi(x))$$

where φ is quantifier-free, possibly with parameters other than the one shown. Since the primitive recursive relations are closed under boolean operations and bounded quantification, the formula above is equivalent, in *PRA*, to a universal one. This fact can be used to show that *PRA* has a universal set of axioms. By Herbrand's theorem, it does not matter whether one takes the underlying logic to be first-order logic, or just the quantifier-free fragment: if the first-order version of *PRA* proves $\forall x\ \exists y\ \varphi(x, y)$ for φ quantifier-free, then there is a function symbol f and a propositional proof of $\varphi(x, f(x))$ from substitution instances of the universal axioms and axioms of equality.

The *finite types* are generated inductively as follows: N is a finite type (denoting the natural numbers, in the intended interpretation); and if σ and τ are types, so are $\sigma \times \tau$ and $\sigma \rightarrow \tau$ (denoting the cross product of σ and τ and the set of functions from σ to τ, respectively, in the full set-theoretic interpretation). I will take the simply-typed lambda calculus to have variables of all finite types, and constants denoting pairing functions, $\langle x, y\rangle$, and projections, $(z)_0$ and $(z)_1$, at all types. The set of lambda terms is further closed under lambda abstraction, denoted $\lambda x\ t$, and application, denoted $t(s)$. I will identify terms that differ up to a renaming of their free variables. If t and s are terms and x is a variable of the appropriate type, then $t[s/x]$ denotes the result of substituting s for x in t, renaming bound variables if necessary. If I introduce a term as $t[x]$, then $t[s]$ abbreviates $t[s/x]$. I will write $t(s_1, \ldots, s_k)$ for $((t(s_1))(s_2)) \ldots (s_k)$, and $\sigma, \tau \rightarrow \rho$ instead of $\sigma \rightarrow (\tau \rightarrow \rho)$. N is sometimes called *type* 0, and a function of type $\mathrm{N}, \ldots, \mathrm{N} \rightarrow \mathrm{N}$ is said to be of *type* 1.

One obtains a higher-type extension of primitive recursive arithmetic as follows. Start with a many sorted version of first-order predicate logic with a sort for each finite type, and an equality relation $=$ at type N only. The terms are the terms of the simply-typed lambda calculus with the extra constants described below. The theory includes equality axioms corresponding to β-reduction; since we only have type N equality, these have to be expressed as schemata:

- $r[(\lambda x.\ t)(s)] = r[t[s/x]]$
- $r[(\langle x, y\rangle)_0] = r[x]$
- $r[(\langle x, y\rangle)_1] = r[y]$,

where in each case $r[z]$ is a term of type N, with z is a variable of the appropriate type. Add a constant 0 of type N and a constant S of type N $\to$ N, with axioms

- $\neg S(x) = 0$
- $S(x) = S(y) \to x = y$.

Then add a constant symbol R of type N, (N, N $\to$ N), N $\to$ N. The idea is that $R(a, f)$ is the function defined by primitive recursion from a and f; hence we have the defining axioms

- $R(a, f, 0) = a$
- $R(a, f, S(x)) = f(x, R(a, f, x))$.

For each type σ add a constant $Cond_\sigma : \mathrm{N}, \sigma, \sigma \to \sigma$ with defining axioms

- $r[Cond(0, x, y)] = r[x]$
- $r[Cond(S(z), x, y)] = r[y]$

for type N terms $r[z]$ with z of type σ. Finally, add a schema of quantifier-free induction, similar to the one for *PRA*. Call the resulting theory PRA^ω.

Using the recursor, R, one can define all the primitive recursive functions. If we identify function symbols of *PRA* with their definitions in PRA^ω, *PRA* is included in PRA^ω. Conversely, we have the following:

THEOREM 2.1. *PRA^ω is a conservative extension of PRA.*

A proof is sketched in [6, Section 5.1]; a similar proof, in the context of polynomial-time computable arithmetic, is found in [15].

I will now describe a nonstandard version of PRA^ω, which I will denote $NPRA^\omega$. First, add a new predicate symbol $st(x)$ to the language, with argument ranging over the natural numbers, and a new constant ω of type N. The predicate st is intended to denote the "standard" natural numbers, while ω is intended to denote a nonstandard natural number. Quantifiers ranging over the standard numbers are obtained by defining $\forall^{st} x\ \varphi$ to be $\forall x\ (st(x) \to \varphi)$ and $\exists^{st} x\ \varphi$ to be $\exists x\ (st(x) \wedge \varphi)$. A formula φ is said to be *internal* if it does not involve st, and *external* otherwise.

To obtain $NPRA^\omega$, add the following axioms to PRA^ω:

- $\neg st(\omega)$
- $st(x) \wedge y < x \to st(y)$
- $st(x_1) \wedge \cdots \wedge st(x_k) \to st(f(x_1, \ldots, x_k))$, for each type 1 term f with no free variables and no occurence of ω.

In particular, the last axiom schema implies that the standard part of the universe is closed under the primitive recursive functions. In addition, add the following schema of $\forall$-transfer without parameters:

- $\forall^{st} \vec{x}\ \psi(\vec{x}) \rightarrow \forall \vec{x}\ \psi(\vec{x})$

where ψ is a quantifier-free internal formula that does not involve ω, in which the only free variables are the type N variables shown.

THEOREM 2.2. *Suppose $NPRA^{\omega}$ proves $\forall^{st} x\ \exists y\ \varphi(x, y)$, where φ is quantifier-free in the language of PRA with the free variables shown. Then PRA^{ω} proves $\forall x\ \exists y\ \varphi(x, y)$, and hence PRA proves it as well.*

Since $\forall^{st} x\ \exists y\ \varphi(x, y)$ is implied by both $\forall x\ \exists y\ \varphi(x, y)$ and $\forall^{st} x\ \exists^{st} y\ \varphi(x, y)$, we have:

COROLLARY 2.3. *$NPRA^{\omega}$ is a Π_2 conservative extension of PRA. Also, if φ is quantifier-free in the language of PRA and $NPRA^{\omega}$ proves $\forall^{st} x\ \exists^{st} y\ \varphi(x, y)$, then PRA proves $\forall x\ \exists y\ \varphi(x, y)$.*

The second part of the corollary indicates a general pattern of reasoning in nonstandard arithmetic, whereby one uses nonstandard numbers to prove theorems about the standard ones.

Let *NPRA* be *PRA* together with the restriction of the axioms above to the smaller language. Since *NPRA* is included in $NPRA^{\omega}$, we have the following:

COROLLARY 2.4. *NPRA is conservative over PRA, in the sense of Theorem* 2.2.

Corollary 2.4 has an easy model-theoretic proof, as follows. Suppose *PRA* does not prove $\forall x\ \exists y\ \varphi(x, y)$. Let L be the language of *PRA*, and let c, d, and ω be new constants. Let T be the set of sentences containing all the following:

- The axioms of *PRA*
- $\forall y\ \neg\varphi(c, y)$
- $d > c$
- $\exists \vec{y}\ \psi(\vec{y}) \rightarrow \exists \vec{y} < d\ \psi(\vec{y})$, for each quantifier-free formula ψ of L with only the free variables shown
- $\omega > t$, for each closed term t of $L + d$.

Every finite subset of T is consistent, since in any model of *PRA* satisfying $\{\exists x\ \forall y\ \neg\varphi(x, y)\}$ we can choose an interpretation of c satisfying $\forall y\ \neg\varphi(c, y)$, an interpretation of d greater than finitely many witnesses for formulae of L of the form $\exists \vec{y}\ \psi(\vec{y})$, and an interpretation of ω greater than the denotation of finitely many terms t involving only d. By compactness, let $\mathcal{M}$ be a model of T. Let S be the set of elements of the universe of $\mathcal{M}$ bounded by a closed term involving only the constant d, i.e.

$$S = \left\{a \in |\mathcal{M}| \mid \text{for some closed term } t \text{ of } L + d,\ a < t^{\mathcal{M}}\right\}.$$

The reader can check that $\mathcal{M}$ becomes a model of *NPRA* satisfying $st(c)$ and $\forall y\ \neg\varphi(c, y)$ when one uses S to interpret the predicate *st*.

A straightforward modification of this argument provides a proof of Theorem 2.2. And, in fact, the argument is much more general, since it relies on very few of the specific features of PRA. In the next section, I will present another proof of Theorem 2.2, by giving an interpretation of $NPRA^\omega$ in PRA^ω. Such an interpretation is interesting in its own right, since it yields an explicit translation, with a polynomial bound on the increase in proof length. In addition, it tells us that Theorem 2.2 can be proved in weak fragments of arithmetic. We will see in Section 4 that the interpretation is almost as general as the model-theoretic argument, and so both are widely applicable.

In comparison with other nonstandard theories, the nonstandard axioms above are fairly weak. I will discuss strengthenings briefly in Sections 4 and 6. But Section 5 suggests that the axioms above are already sufficient to formalize an interesting portion of real analysis.

§3. The interpretation. The interpretation of $NPRA^\omega$ in PRA^ω uses a forcing argument, described entirely in the language of PRA^ω. For similar forcing arguments, see [2, 3, 5, 7, 9].

Let L denote the (typed) language of PRA^ω, and L^{st} denote the language of $NPRA^\omega$, i.e. L together with an extra constant, ω, and a new predicate, $st(x)$. Our first step is to translate terms of L^{st} to terms of L. Choose a type N variable, ω, in the language of L, corresponding to the constant, ω, of L^{st}. Also, assign to each variable x of type σ in L^{st} a variable $\tilde{x}$ of type $\mathrm{N} \to \sigma$ in L. Finally, if $t[x_1, \ldots, x_k]$ is a term of L with the free variables shown, let $\widehat{t}$ denote the term $t[\tilde{x}_1(\omega), \ldots, \tilde{x}_k(\omega)]$ of L, where the constant ω of L^{st} is also replaced by the corresponding variable of L.

The idea is that we are taking elements of the universe of L^{st} to be named by terms of L that depend on a "generic" element, ω. It is not hard to check that the axioms of β-reduction are preserved by the translation.

Conditions of the forcing relation are ternary relations, considered as elements of type $\mathrm{N}, \mathrm{N}, \mathrm{N} \to \mathrm{N}$. Intuitively, a condition p is supposed to represent the assertion $\forall^{st} u \; \forall v \;\; p(u, v, \omega)$, where ω is the generic nonstandard element. If p and q are conditions, define $p \preceq q$ to be the formula $\forall u, v, \omega \; (p(u, v, \omega) \to q(u, v, \omega))$, read "$p$ is stronger than (or equivalent to) q." Note that if p and q are conditions, then their conjunction, $p \wedge q$, satisfies $p \wedge q \preceq p$ and $p \wedge q \preceq q$. Sometimes, if p is a condition and A is another ternary relation, I will write $p \wedge \forall^{st} u \; \forall v \;\; A(u, v, \omega)$ instead of $p \wedge A$. This is nothing more than a useful convention that will keep us mindful of the informal interpretation of the conditions.

We are now ready to define a relation $p \Vdash \varphi$ between conditions p and formulae φ of L^{st}. It will be convenient to take the logical connectives to be $\forall$, $\wedge$, $\to$, $\bot$. With this choice of connectives, $\neg\varphi$ abbreviates $\varphi \to \bot$, $\exists x \; \varphi$ abbreviates $\neg \forall x \; \neg\varphi$, and $\varphi \vee \psi$ abbreviates $\neg(\neg\varphi \wedge \neg\psi)$. The forcing relation is defined inductively, as follows:

1. $p \Vdash \bot \equiv \exists z\ \forall \omega\ \neg \forall u < z\ \forall v\ p(u, v, \omega)$.
2. $p \Vdash t_1 = t_2 \equiv \exists z\ \forall \omega\ (\forall u < z\ \forall v\ p(u, v, \omega) \to \widehat{t_1} = \widehat{t_2})$.
3. $p \Vdash t_1 < t_2 \equiv \exists z\ \forall \omega\ (\forall u < z\ \forall v\ p(u, v, \omega) \to \widehat{t_1} < \widehat{t_2})$.
4. $p \Vdash st(t) \equiv \exists z\ \forall \omega\ (\forall u < z\ \forall v\ p(u, v, \omega) \to \widehat{t} < z)$.
5. $p \Vdash \varphi \to \psi \equiv \forall q \preceq p\ (q \Vdash \varphi \to q \Vdash \psi)$.
6. $p \Vdash \varphi \wedge \psi \equiv (p \Vdash \varphi) \wedge (p \Vdash \psi)$.
7. $p \Vdash \forall x\ \varphi \equiv \forall \tilde{x}\ (p \Vdash \varphi)$

If φ is a formula in the language L^{st} with free variables $x_1, \dots, x_k$, then $p \Vdash \varphi$ is a formula in the language L with free variables $\omega, \tilde{x}_1, \dots, \tilde{x}_k$, as well as p. Notice that we are allowing that some conditions force $\bot$. In the definition of $p \Vdash \varphi \to \psi$, the quantifier $\forall q \preceq p$ ranges over conditions. It is not difficult to show that $p \Vdash \varphi \to \psi$ is equivalent to $\forall q\ (q \Vdash \varphi \to p \wedge q \Vdash \psi)$; I will use both formulations of $p \Vdash \varphi \to \psi$ below. Define $\Vdash \varphi$ to be $\forall p\ (p \Vdash \varphi)$, read "$\varphi$ is forced."

The following informal considerations may help explain the motivation behind the definition of forcing at the atomic clauses. Think of a condition p as representing an infinite set of sentences,

$$\{\forall v\ p(0, v, \omega), \forall v\ p(1, v, \omega), \forall v\ p(2, v, \omega), \dots\}.$$

If we call this set S_p then clause 2, for example, asserts that p forces $t_1 = t_2$ if and only if $\widehat{t_1} = \widehat{t_2}$ is a consequence of a finite subset of S_p.

The proofs of the next five lemmata are routine and standard. (See, for example, [3, 9] for a little more detail.)

Lemma 3.1. *Suppose t and s are terms of L^{st}, $r[z]$ is a type* N *term of PRA^ω, and z has the same type as t. Then PRA^ω proves*

$$r\left[\widehat{t}\left[\lambda\omega\ \widehat{s}/\tilde{x}\right]\right] = r\left[\widehat{t[s/x]}\right].$$

Proof. By induction on t. Informally, in the base case where t is x, we have

$$\widehat{t}\left[\lambda\omega\ \widehat{s}/\tilde{x}\right] = \tilde{x}(\omega)\left[\lambda\omega\ \widehat{s}/\tilde{x}\right] = (\lambda\omega\ \widehat{s})(\omega) = \widehat{s} = \widehat{t[s/x]}.$$

The other cases are easy. ⊣

Lemma 3.2 (substitution). *For each formula φ and term s in the language of L^{st}, PRA^ω proves $p \Vdash \varphi[s/x] \leftrightarrow (p \Vdash \varphi)[\lambda\omega\ \widehat{s}/\tilde{x}]$.*

Proof. By induction on φ. Lemma 3.1 takes care of the base cases. ⊣

Lemma 3.3 (monotonicity). *For each formula φ of L^{st}, PRA^ω proves $p \Vdash \varphi \wedge q \preceq p \to q \Vdash \varphi$.*

Proof. Induction on φ. ⊣

Lemma 3.4. *For each formula φ in the language of L^{st}, PRA^ω proves $\Vdash (\bot \to \varphi)$.*

Proof. Induction on φ. ⊣

and hence

$$\exists z\ \forall\omega\ (\forall u < z\ \forall v\ q(u,v,\omega) \to \widehat{\eta}).$$

By the inductive hypothesis, this is equivalent to $q \Vdash \eta$. ⊣

LEMMA 3.10. *If φ is an axiom of PRA^ω, then PRA^ω proves $\Vdash \varphi$.*

PROOF. All the axioms of PRA^ω are universal, which is to say, they are of the form $\forall x_1,\dots,x_k\ \varphi(x_1,\dots,x_k)$, where φ is quantifier-free and internal. By Lemma 3.9, $\Vdash \forall\vec{x}\ \varphi$ is equivalent to $\forall\tilde{x}_1,\dots,\tilde{x}_k\ \forall\omega\ \varphi(\tilde{x}_1(\omega),\dots,\tilde{x}_k(\omega))$. The axioms corresponding to β-reduction are easy to verify, and otherwise, the translation of each axiom follows immediately from the untranslated version. ⊣

LEMMA 3.11. *For each constant term f of type $\mathrm{N}^k \to \mathrm{N}$ of L^{st} not involving ω, PRA^ω proves $\Vdash \forall x_1,\dots,x_k\ (st(x_1)\wedge\cdots\wedge st(x_k) \to f(x_1,\dots,x_k))$.*

PROOF. The key point is that if f is as in the hypothesis, it does not depend on ω. Argue in PRA^ω. Suppose $p \Vdash (st(x_1)\wedge\ldots\wedge st(x_k))$, i.e. $\exists z\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to (\tilde{x}_1(\omega) < z \wedge\ldots\wedge \tilde{x}_k(\omega) < z))$. Letting $z' = \max_{\vec{v}<z} \widehat{f}(v_1,\dots,v_k)$, we have $\exists z,z'\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to \widehat{f}(\tilde{x}_1(\omega),\dots,\tilde{x}_k(\omega)) < z')$, which implies $p \Vdash st(f(x_1,\dots,x_k))$. ⊣

LEMMA 3.12. *PRA^ω proves $\Vdash \forall x,y\ (st(x)\wedge y < x \to st(y))$.*

PROOF. Argue in PRA^ω. Suppose $p \Vdash st(x)$ and $p \Vdash y < x$. Then $\exists z\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to \tilde{x}(\omega) < z)$ and $\exists z\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to \tilde{y}(\omega) < \tilde{x}(\omega))$. Picking z to be the maximum of any two witnesses to these statements, we have $\exists z\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to \tilde{y}(\omega) < z)$, which is $p \Vdash st(y)$. ⊣

We have dealt with all the axioms except for $\forall$-transfer and $\neg st(\omega)$. The next lemma deals with the former.

LEMMA 3.13. *If $\varphi(\vec{x})$ is any quantifier-free formula of L with only the type N variables shown, PRA^ω proves $\Vdash \forall^{st}\vec{x}\ \varphi(\vec{x}) \to \forall\vec{x}\ \varphi(\vec{x})$.*

PROOF. For notational simplicity, let us assume that $\vec{x}$ is a single variable. Argue in PRA^ω. Suppose $p \Vdash \forall^{st}x\ \varphi(x)$. By Lemma 3.6, we have $\forall w\ ((p \Vdash \varphi(x))[\lambda\omega\ w/\tilde{x}])$. Since $\widehat{\varphi(x)}[\lambda\omega\ w/\tilde{x}]$ is equivalent to $\varphi(w)$, by Lemma 3.9 we have

$$\forall w\ \exists z\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to \varphi(w)).$$

Since φ does not depend on ω or z, this is equivalent to

$$\exists z\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to \forall w\ \varphi(w)),$$

which in turn implies

$$\forall\tilde{x}\ \exists z\ \forall\omega\ (\forall u < z\ \forall v\ p(u,v,\omega) \to \varphi(\tilde{x}(\omega))).$$

But the last formula is equivalent to $p \Vdash \forall x\ \varphi(x)$, which is what we want. ⊣

LEMMA 3.14. *Suppose* φ *is any formula of* L^{st}, *and* $NPRA^\omega$ *proves* φ. *Then* PRA^ω *proves* $\forall^{st} u\ (\omega \not< u) \Vdash \varphi$.

PROOF. By Lemma 3.6, we have $\forall^{st} u\ (\omega \not< u) \Vdash \neg st(\omega)$, and we have shown that all the other axioms of $NPRA^\omega$ are forced. ⊣

We are now only one lemma away from the proof of the main theorem.

LEMMA 3.15. *Suppose* $\varphi(y, x_1, \ldots, x_k)$ *is a quantifier-free internal formula of* L^{st} *with the free variables shown, and* y *is of type* N. *Then* PRA^ω *proves* $\forall v\ \varphi(v, \tilde{x}_1(\omega), \ldots \tilde{x}_k(\omega)) \Vdash \forall y\ \varphi(y, x_1, \ldots, x_k)$.

PROOF. Unwinding the definition and using Lemma 3.9, we see that we need to show that PRA^ω proves

$$\forall \tilde{y}\ \exists z\ \forall \omega\ (\forall v\ \varphi(v, \tilde{x}_1(\omega), \ldots, \tilde{x}_k(\omega)) \rightarrow \varphi(\tilde{y}(\omega), \tilde{x}_1(\omega), \ldots, \tilde{x}_k(\omega))).$$

But this is immediate. ⊣

PROOF OF THEOREM 2.2. Suppose $NPRA^\omega$ proves $\forall^{st} x\ \exists y\ \varphi(x, y)$ with φ quantifier-free in the language of L, and argue in PRA^ω. By Lemma 3.14, we have

$$\forall^{st} u\ (\omega \not< u) \Vdash \forall^{st} x\ \exists y\ \varphi(x, y).$$

Let w be arbitrary. Since $(p \Vdash st(x))[\lambda\omega\ w/\tilde{x}]$, we have

$$(\forall^{st} u\ (\omega \not< u) \Vdash \exists y\ \varphi(x, y))[\lambda\omega\ w/\tilde{x}].$$

Keep in mind that $\exists y\ \varphi(x, y)$ abbreviates $\neg\forall y\ \neg\varphi(x, y)$. By the previous lemma, $\forall v\ \neg\varphi(\tilde{x}(\omega), v) \Vdash \forall y\ \neg\varphi(x, y)$, so we have

$$(\forall^{st} u\ (\omega \not< u) \wedge \forall v\ \neg\varphi(\tilde{x}(\omega), v) \Vdash \bot)[\lambda\omega\ w/\tilde{x}],$$

which expands to

$$\exists z\ \forall \omega\ (\forall u < z\ \forall v\ (\omega \not< u \wedge \forall v\ \neg\varphi(w, v)) \rightarrow \bot).$$

This is classically equivalent to

$$\exists z\ \forall \omega\ (\exists u < z\ (\omega < u) \vee \exists v\ \varphi(w, v)).$$

Given a z witnessing this statement, choose $\omega = z$. Then we have $\forall u < z\ (\omega \not< u)$, and hence $\exists v\ \varphi(w, v)$. Since w was arbitrary, we have $\forall w\ \exists v\ \varphi(w, v)$, as desired. ⊣

§4. Weak theories of nonstandard arithmetic. In this section, I will discuss variations of Theorem 2.2, and some applications. To start with, there are a number of features of Theorem 2.2 and its proofs that are worth noting.

The first has to do with the treatment of equality. The theories PRA^ω and $NPRA^\omega$ were presented with only equality at type N as a basic relation. Of course, one can define equality at higher types extensionally; for example, if f and g are of type $N \rightarrow N$ one can take $f = g$ to be $\forall x\ (f(x) = g(x))$. Doing so does not guarantee that the usual equality axioms, $f = g \rightarrow \varphi(f) = \varphi(g)$, follow. But, using a method due to Luckhardt, one can interpret a fully extensional version of $NPRA^\omega$ in our intensional version, by relativizing all quantifiers and variables to the "hereditarily extensional objects." This interpretation preserves Π_2 formulae (as well as $\forall^{st} \exists^{st}$ formulae, etc.). So Theorem 2.2 extends to extensional versions of $NPRA^\omega$ as well. For a discussion of some of the issues related to various treatments of equality, see [6, Section 3.1], [45, Section 3.1], and [15, Section 7].

Second, most of the higher types were not used by the intepretation in an essential way. It suffices to have a theory in which the types are closed under the operation $\sigma \mapsto (\mathrm{N} \to \sigma)$, so, for example, the interpretation works just as well for second-order versions of $NPRA$ and PRA, associating k-ary function variables of the first theory to $(k+1)$-ary function variables of the second.

Finally, very little reference was made to the specifics of PRA^ω itself. In the interpretation, only the following features came into play:

1. PRA^ω proves that $<$ is transitive and anti-reflexive, and satisfies the sentence $\forall x, y\ \exists z\ (x < z \wedge y < z)$.
2. PRA^ω has a universal set of axioms.
3. If $\varphi(x, y, z)$ is a quantifier-free formula, possibly with free variables shown, PRA^ω proves
$$\exists R\ \forall x, y, z\ (\varphi(x, y, z) \leftrightarrow R(x, y, z)),$$
where R ranges over a suitable representation of ternary relations.
4. PRA^ω proves that the ternary relations are closed under conjunction.
5. If f is a closed type 1 term, then PRA^ω proves $\forall z\ \exists w\ \forall \vec{x} < z\ (f(\vec{x}) < w)$. (This was used in the proof of Lemma 3.11.)

In fact, most of the proofs in the previous section required only the intuitionistic fragment of PRA^ω. The proof of Lemma 3.13, which showed that the $\forall$-transfer schema without parameters is forced, used classical logic. But if one is willing to give up transfer, then only the final proof of Theorem 2.2 requires an inference that is not strictly intuitionistic, in the form of Markov's principle for quantifier-free formulae; and a slight rewriting of the forcing relation, along the lines of [3], can be used to render the argument entirely intuitionistic. On the other hand, if one has no qualms about the use of classical logic, the presentation of the forcing relation can be simplified; see Appendix B below.

In sum, both the model-theoretic argument sketched at the end of Section 2 and the syntactic interpretation of Section 3 generalize considerably. For example, let PV be Cook's theory of polynomial-time computable functions, and let PV^ω be a corresponding higher-type generalization (i.e. the theory called PV^ω in [15], but with induction restricted to quantifier-free formulae; see also [6, Section 5.2]). Let NPV^ω be the nonstandard version obtained by adding the nonstandard axioms of Section 2. Then we have

THEOREM 4.1. *NPV^ω is conservative over PV^ω and PV, in the sense of Theorem* 2.2.

Similarly, let *ERA* denote elementary recursive arithmetic, obtained by adding $+$, $\times$, and x^y to *PRA* but restricting the recursions to those that can be bounded by a term (see e.g. [4, 39, 40]). *ERA* is a conservative extension of the theory alternatively known as *EFA*, for "elementary function arithmetic", or $I\Delta_0(exp)$ (see [23]). Let ERA^ω be the natural higher-type version of *ERA* (similar to the theory G_3A^ω of [28, Section 2.2], but without the additional universal sentences in clause 9), and let $NERA^\omega$ be the nonstandard version of ERA^ω. Then we have

THEOREM 4.2. *$NERA^\omega$ is conservative over ERA^ω and ERA, in the sense of Theorem* 2.2.

Similar versions of Theorem 2.2 hold, for example, for the theory denoted $T + (\mu)$ in [6], and for the theories G_nA^ω of [28].

The nonstandard axioms can have interesting consequences for the standard numbers. Recall that a formula is *bounded*, or Δ_0, if all its quantifiers are bounded, and Σ_1 if it is of the form $\exists\vec{x}\ \varphi$, where φ is bounded. Consider the collection principle, $B\Sigma_1$:

$$\forall x < z\ \exists y\ \varphi(x, y) \rightarrow \exists w\ \forall x < z\ \exists y < w\ \varphi(x, y)$$

where φ is Σ_1. Let $B\Sigma_1^{st}$ denote the relativization of $B\Sigma_1$ to the standard numbers, where z is assumed to be standard. The following proposition still holds even if φ has additional parameters that are not necessarily standard.

PROPOSITION 4.3. *$NERA^\omega$ and $NPRA^\omega$ prove $B\Sigma_1^{st}$.*

PROOF. By pairing existential quantifiers, we may assume φ is Δ_0. Argue in $NERA^\omega$ or $NPRA^\omega$. Suppose z is standard and $\forall x < z\ \exists^{st} y\ \varphi(x, y)$. Then for any nonstandard number w, we have $\forall x < z\ \exists y < w\ \varphi(x, y)$. Since, in $NERA^\omega$ and $NPRA^\omega$, every bounded formula is equivalent to a quantifier-free (even atomic) one, by induction there is a least w such that this last formula is satisfied. Since the nonstandard numbers are closed under predecessor, this least w is standard. ⊣

Another interesting fact is that we can interpret the theory WKL_0^* of Simpson and Smith [37, 38]. This theory is equivalent to a second-order version of *ERA* with set variables, together with a recursive comprehension axiom, (*RCA*), and a weak version of König's lemma, (*WKL*), which asserts that every infinite tree on $\{0, 1\}$ has an infinite path. For details, see [37, 38].

THEOREM 4.4. *WKL_0^* is conservative over ERA for Π_2 formulae.*

PROOF. One can interpret WKL_0^* in *NERA*, interpreting the first-order universe as the standard numbers of *NERA*, and interpreting the second-order universe as the standard parts of nonstandard finite sets of *NERA*. Here we are using the fact that in *NERA*, one can code finite sets as natural numbers;

note that if $\mathcal{M}$ is a model of *NERA* and S is a set coded in $\mathcal{M}$, then the intersection of S with the standard numbers of $\mathcal{M}$ may be unbounded.

Lemma IV.4.4 of [37] shows that (WKL) and (RCA) follow from a single schema of Σ_1 separation:

$$\forall x\ \neg(\exists y\ \varphi(x,y) \wedge \exists z\ \psi(x,z)) \\ \rightarrow \exists S\ \forall x\ ((\exists y\ \varphi(x,y) \rightarrow x \in S) \wedge (\exists z\ \psi(x,z) \rightarrow x \notin S)),$$

where φ and ψ are Δ_0. To see that this holds in the interpretation, argue in *NERA*. Suppose for every standard x we have $\neg(\exists^{st} y\ \varphi(x,y) \wedge \exists^{st} z\ \psi(x,z))$. Let S be the finite set

$$S = \{x < \omega \mid \exists y < \omega\ (\varphi(x,y) \wedge \forall z < y\ \neg\psi(x,z))\}.$$

It is not hard check that for each standard x, we have

$$(\exists^{st} y\ \varphi(x,y) \rightarrow x \in S) \wedge (\exists^{st} z\ \psi(x,z) \rightarrow x \notin S),$$

as required. ⊣

The results of [38] are more general; for example, the first-order consequences of WKL_0^* are exactly those of $ERA + B\Sigma_1$.

In Section 6, we will see that, at least for the case of ERA^ω, the transfer principles and induction in the system are close to optimal. When it comes to *PRA*, however, it seems worth mentioning another conservation result that can be obtained by the entirely different methods. Let $(I\Sigma_1^{st})$ denote the relativization of the schema of Σ_1 induction to the standard numbers. (Here too it does not hurt if we allow nonstandard parameters.) Let $NPRA'$ consist of *NPRA* without the $\forall$-transfer axiom, together with $(I\Sigma_1^{st})$. Then we have the following:

THEOREM 4.5. *If $NPRA'$ proves $\forall^{st} x\ \exists^{st} y\ \varphi(x,y)$ with φ quantifier-free in the language of PRA, then PRA proves $\forall x\ \exists y\ \varphi(x,y)$.*

PROOF. The corresponding theorem for an intuitionistic version of $NPRA'$ is proved in [7, Theorem 4.4]. By [7, Lemma 5.1], this intuitionstic theory proves that Markov's principle for primitive recursive relations holds on the standard numbers. Our $NPRA'$ can therefore be interpreted in the intuitionistic version, using a double-negation translation. ⊣

Section 6 raises the question as to whether or not there is a common refinement of Theorems 2.2 and 4.5. Nonetheless, Theorem 4.5 is strong enough to yield the following celebrated result of Friedman. Here WKL_0 is essentially WKL_0^* together with the schema of Σ_1 induction.

THEOREM 4.6. *WKL_0 is conservative over PRA for Π_2 sentences.*

PROOF. As in the proof of Theorem 4.4, WKL_0 is interpreted in $NPRA'$. ⊣

A further connection between Weak König's lemma and nonstandard analysis is discussed in [42].

The usual field operations on $\mathbb{Q}^*$ lift to make $\mathbb{R}$ an ordered field. Under this lifting, division by 0 can have unusual properties; for example, if p is $1/\omega$ and q is $2/\omega$, then, as real numbers, $p = q = 0$, but $p/q = 1/2$.

Let us pause for a moment to compare this to common developments of nonstandard analysis (as in, say, [22]). In such developments, one typically defines the nonstandard reals, $\mathbb{R}^*$, in which one can embed the standard ones; and any standard function f from $\mathbb{R}$ to $\mathbb{R}$ has a nonstandard extension f^* from $\mathbb{R}^*$ to $\mathbb{R}^*$. In our setup, nonstandard reals would have to be developed as type 1 objects, e.g. as Cauchy sequences of nonstandard rationals; general functions from $\mathbb{R}^* \to \mathbb{R}^*$ would then be type 2. Here I propose to ignore the nonstandard reals entirely. We will see below that though this approach has some quirks, it suffices for the development of parts of real analysis, and it has the advantage that real numbers are represented by type 0 objects.

The following lemma says that one can bound the size of the numerator and denominator in the nonstandard representation of a real number.

LEMMA 5.1. *Let x be an element of $\mathbb{R}$. Then there are $a \in \mathbb{Z}^*$ and $b \in \mathbb{N}^*$ such that $|a| < \omega$, $b < \omega$, and $x =_{\mathbb{R}} a/b$.*

PROOF. If $x =_{\mathbb{Q}^*} 0$, take $a = 0$, $b = 1$. Otherwise, we can assume $x >_{\mathbb{Q}^*} 0$; if $x <_{\mathbb{Q}^*} 0$, apply the argument to $-x$.

Since $x \in \mathbb{R}$, x is bounded by a standard natural number $c = \lceil x \rceil > 0$. So if we let $b = \lfloor (\omega - 1)/c \rfloor$, b is nonstandard as well. We want to find a such that

$$\frac{a}{b} \leq_{\mathbb{Q}^*} x <_{\mathbb{Q}^*} \frac{a+1}{b}$$

so let $a = \lfloor bx \rfloor$. Then $b < \omega$ and $a \leq_{\mathbb{Q}^*} bx \leq \lfloor (\omega - 1)/c \rfloor c \leq \omega - 1 < \omega$. Since $x - a/b <_{\mathbb{Q}^*} 1/b$, we have $x =_{\mathbb{R}} a/b$, as needed. $\dashv$

The proof shows moreover that suitable values of a and b can be computed from x by a type 1 term of ERA^ω. The advantage bestowed by Lemma 5.1 is that certain quantifiers over the real numbers become equivalent to bounded ones. For example, suppose $\varphi(x)$ is a formula which respects equality of reals. Then $\forall x^{\mathbb{R}}\ \varphi(x)$ is equivalent to $\forall a^{\mathbb{N}^*} < \omega,\ b^{\mathbb{N}^*} < \omega\ (b \neq 0 \wedge st(\lceil a/b \rceil) \to \varphi(\pm a/b))$. This last formula is external, since it involves st. But if φ is internal and one wants to quantify over a *bounded* range of real numbers, one can replace $st(\lceil a/b \rceil)$ by an explicit bound, in which case the result is an internal formula. So, for example, if $R(x, y)$ is a relation (i.e. a type 1 term) in $NERA^\omega$ that respects equality on the real numbers and r and s are reals, then $\forall x \in [r, s]\ R(x, y)$ is also equivalent to a relation in $NERA^\omega$.

A function f from $\mathbb{R}$ to $\mathbb{R}$ is *continuous* if it satisfies the usual ε-δ definition of continuity:

$$\forall x^{\mathbb{R}}\ \forall \varepsilon^{\mathbb{R}} > 0\ \exists \delta^{\mathbb{R}} > 0\ \forall y^{\mathbb{R}}\ (|x - y| < \delta \to |f(x) - f(y)| < \varepsilon).$$

In $NERA^\omega$ we have the following surprising fact:

Proposition 5.2. *Every function $f \in \mathbb{R} \to \mathbb{R}$ is continuous.*

Proof. Suppose we are given $f \in \mathbb{R} \to \mathbb{R}$, $x \in \mathbb{R}$, and $\varepsilon \in \mathbb{R}$ with $\varepsilon >_{\mathbb{R}} 0$. It suffices to find a $\delta \in \mathbb{R}$ such that

$$\forall h^{\mathbb{R}} \left(|h| < \delta \to |f(x+h) - f(x)| < \varepsilon\right). \tag{2}$$

Since f respects equality on $\mathbb{R}$, we know that for each nonstandard natural number m,

(3)
$$\forall a^{\mathbb{Z}^*} \in (-\omega, \omega)\, \forall b < \omega \left(\left|a/b\right| <_{\mathbb{Q}^*} 1/m \to \left|f\left(x + a/b\right) - f(x)\right| <_{\mathbb{Q}^*} \varepsilon/2\right).$$

If (3) holds for all $m \geq 1$ let $m = 0$, and otherwise, by induction, let m be the greatest number less than ω such that (3) fails. Let $\delta = 1/(m+1)$. Since m is standard, $\delta >_{\mathbb{R}} 0$.

I claim that this δ satisfies (2). Suppose $|h| <_{\mathbb{R}} \delta$. By Lemma 5.1, $h =_{\mathbb{R}} a/b$ for some $a \in \mathbb{Z}^*$ and $b \in \mathbb{N}^*$ with $|a|, b < \omega$. Then $|a/b| <_{\mathbb{Q}^*} 1/(m+1)$, and so $|f(x + a/b) - f(x)| <_{\mathbb{Q}^*} \varepsilon/2$. Since f is a function on $\mathbb{R}$ and $h =_{\mathbb{R}} a/b$, we have $|f(x+h) - f(x)| \leq_{\mathbb{R}} \varepsilon/2 <_{\mathbb{R}} \varepsilon$. ⊣

The proof above used induction on a bounded formula, and so does not go through in NPV^{ω}. But in NPV^{ω} one can prove the converse, namely, that every function $f \in \mathbb{Q}^* \to \mathbb{Q}^*$ satisfying the continuity condition is in fact a function $f \in \mathbb{R} \to \mathbb{R}$.

At first glance, Proposition 5.2 seems blatantly false. After all, what about the function $f \in \mathbb{Q}^* \to \mathbb{Q}^*$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x \leq_{\mathbb{Q}^*} 0 \\ 1 & \text{otherwise,} \end{cases}$$

which is represented by a term of $NERA^{\omega}$? The problem is that this is not a function from $\mathbb{R}$ to $\mathbb{R}$: for example, $1/\omega =_{\mathbb{R}} 0$ but $f(1/\omega) \neq_{\mathbb{R}} f(0)$. On the other hand, the function $g \in \mathbb{Q}^* \to \mathbb{Q}^*$ defined by

$$g(x) = \begin{cases} 0 & \text{if } x \leq_{\mathbb{R}} 0 \\ 1 & \text{otherwise} \end{cases}$$

is not represented by a term of $NERA^{\omega}$, since $x \leq_{\mathbb{R}} 0$ is external.

Thus, we have a development of analysis which, like Brouwer's, has the property that every well-defined function from $\mathbb{R}$ to $\mathbb{R}$ is continuous. This feature may help illuminate the Brouwerian world-view. What is going on is that in our framework, function variables $f, g, \ldots$ range over internal functions; and, in essence, Proposition 5.2 tells us that any function from $\mathbb{R} \to \mathbb{R}$ defined from an internal function from $\mathbb{Q}^* \to \mathbb{Q}^*$ is continuous. If one is loathe to give up functions like g above, one can extend our theories with function variables ranging over *external* functions, which are not allowed to appear in the induction axioms. Thus, in a sense, Proposition 5.2 is not incompatible with classical mathematics; it only underscores the fact that, in the theory at hand, we have chosen to ignore the additional functions. For many purposes this restriction poses no great loss; for example, the function f above is well-defined on any interval of $\mathbb{R}$ that does not contain 0.

The examples that follow provide evidence that our framework allows a smooth development of elementary calculus.

THEOREM 5.3. *If $f \in [0, 1] \to \mathbb{R}$, then f is uniformly continuous.*

PROOF. The proof is similar to that of Proposition 5.2 above. ⊣

THEOREM 5.4 (Intermediate value theorem). *Suppose $f \in [0, 1] \to \mathbb{R}$, $f(0) = -1$, and $f(1) = 1$. Then there is an $x \in [0, 1]$ such that $f(x) = 0$.*

PROOF. Considering f as a function on $\mathbb{Q}^*$, let

$$j = \max\left\{i < \omega \mid f\left(i/\omega\right) <_{\mathbb{Q}^*} 0\right\}$$

and let $x = j/\omega$. Since $j/\omega \sim (j+1)/\omega$, we have

$$f((j+1)/\omega) =_{\mathbb{R}} f(j/\omega) \leq_{\mathbb{R}} 0 \leq_{\mathbb{R}} f((j+1)/\omega)$$

and so $f(x) =_{\mathbb{R}} 0$. ⊣

THEOREM 5.5 (Extreme value theorem). *If $f \in [0, 1] \to \mathbb{R}$, then f attains a maximum value.*

PROOF. Again considering f as a function on $\mathbb{Q}^*$, let

$$y = \max_{0 \leq i \leq \omega} f(i/\omega),$$

let $x = j/\omega$ satisfy $f(x) =_{\mathbb{Q}^*} y$. That y is a maximum is guaranteed by the fact that for any $x' \in [0, 1]$, there is an i such that $x' \sim i/\omega$. ⊣

Turning to differentiation, if $f \in \mathbb{R} \to \mathbb{R}$ and $x, y \in \mathbb{R}$, say $f'(x) = y$ if

$$\forall \varepsilon^{\mathbb{R}} > 0\ \exists \delta^{\mathbb{R}} > 0\ \forall h^{\mathbb{R}} \left(0 \neq_{\mathbb{R}} |h| < \delta \to \left|\frac{f(x+h) - f(x)}{h} - y\right| < \varepsilon\right).$$

This is not the strongest condition one can imagine, since it says nothing about the behavior of f at nonzero infinitesimals. For example, it is possible that $f'(0) = 0$ while, as a function from $\mathbb{Q}^*$ to $\mathbb{Q}^*$, $f(x)$ oscillates between $-x$ and x on an infinitesimal neighborhood of 0. Say that *$f'(x)$ is strongly equal to y* if the formula above holds with $\neq_{\mathbb{R}}$ replaced by $\neq_{\mathbb{Q}}$.

PROPOSITION 5.6. *Let $f \in \mathbb{R} \to \mathbb{R}$, $x, y \in \mathbb{R}$. Then the following are equivalent*:

1. *$f'(x)$ is strongly equal to y.*
2. *For every infinitesimal $h \neq_{\mathbb{Q}^*} 0$, $\frac{f(x+h)-f(x)}{h} \sim y$.*

If $f'(x)$ is strongly equal to y, then $f'(x) = y$.

Proof. The last claim is obvious. For the implication $1 \rightarrow 2$, suppose $f'(x)$ is strongly equal to y and let h be a nonzero infinitesimal element of $\mathbb{Q}^*$. Then for every $\delta^{\mathbb{R}} > 0$, $|h| <_{\mathbb{R}} \delta$. So for every standard n, $|(f(x+h) - f(x))/h - y| < 1/n$. This implies that $(f(x+h) - f(x))/h - y$ is infinitesimal.

The proof that $2 \rightarrow 1$ is very similar to that of Proposition 5.2. ⊣

Corollary 5.7. *Suppose k is standard, and $f(x) = x^k$. Then for every x, $f'(x) = kx^{k-1}$.*

Proof. Suppose h is infinitesimal. Calculating, we have

$$\frac{(x+h)^k - x^k}{h} = kx^{k-1} + h\left[\sum_{i=2}^{k} \binom{k}{i} x^{k-i} h^{i-2}\right].$$

Using the facts that k and $\lceil x \rceil$ are standard, it is not hard to show that the expression in brackets is standard, and so its product with h is infinitesimal. ⊣

One can continue, for example, by defining functions like e^x, $\sin x$, and $\cos x$ using nonstandard finite segments of their Taylor series expansions, and then deriving their basic properties. For another example, there is an easy proof of the Cauchy-Peano theorem on the existence of solutions to differential equations, as described in [42]. There does not seem to be any bar to developing integral calculus in $NERA^\omega$ in a similar manner.

§**6. Notes and questions.** This paper is a modest contribution to the study of weak theories of nonstandard arithmetic, and there are a number of questions and issues that need to be further explored. The questions discussed in this section fall into two groups: the first has to do with the metamathematical properties of the formal theories under consideration, and the second has to do with their utility with respect to the formal analysis of mathematics.

As far as the theories go, one would like to know the extent to which they are optimal, and whether or not they can be strengthened with additional principles of induction, transfer, and so on, while maintaining Π_2 conservativity. For example, we might want to strengthen the $\forall$-transfer axiom of Section 2 by allowing standard parameters: $\forall^{st} \vec{y}\, (\forall^{st} \vec{x}\, \varphi(\vec{x}, \vec{y}) \rightarrow \forall \vec{x}\, \varphi(\vec{x}, \vec{y}))$, where φ is an internal Δ_0 formula that does not involve ω. The following shows that we cannot even add this mild strengthening to $NERA^\omega$ without violating Π_2 conservativity:

Proposition 6.1. *Over $NERA^\omega$, $\forall$-transfer with parameters implies Σ_1 induction relative to the standard numbers, i.e. the schema*

$$\forall^{st} \vec{y}\, (\psi(0, \vec{y}) \wedge \forall^{st} u\, (\psi(u, \vec{y}) \rightarrow \psi(u+1, \vec{y})) \rightarrow \forall^{st} u\, \psi(u, \vec{y})),$$

for $\psi(u, \vec{y})$ of the form $\exists^{st} x\, \varphi(u, x, \vec{y})$, where φ is a Δ_0 formula in the langauge of ERA^ω.

PROOF. $\forall$-transfer with parameters implies that if $\varphi(u, x, \vec{y})$ is as above and u and $\vec{y}$ are assumed to be standard, then $\exists^{st} x\ \varphi(u, x, \vec{y})$ is equivalent to $\exists x < \omega\ \varphi(u, x, \vec{y})$. By induction for bounded formulae in $NERA^{\omega}$, if $\exists x < \omega\ \varphi(u, x, \vec{y})$ fails for some u, there is a least such u; and is least u has to be nonstandard. ⊣

This leaves open the question as to whether one can improve the conservation result for $NPRA^{\omega}$, using either the methods presented here or in [7]. How much transfer can one add? Can one add the unrelativized version of Σ_1 induction for formulae in the original language? In trying to strengthen the conservation result, one might make use of the fact that one can add Σ_1 induction, and even Π_2 collection, to PRA^{ω} without destroying Π_2 conservativity. Formalized or internalized versions of the various model-theoretic constructions presented in [23, 24, 5] may also be useful in this regard.

Similarly, one can extend PV with Σ_1^b induction, yielding, essentially, Buss' theory S_2^1 (for various formulations, see [10, 12, 15, 19]). And one can extend S_2^1 with either a weak form of collection for arbitrary bounded formulae (see [11], or [20] for a simpler model-theoretic proof) or a stronger form of collection for Σ_2^b formulae (see [12]). Can either of these results or the associated model-theoretic constructions be used to strengthen the theory NPV^{ω}? In particular, can one obtain a strengthening of NPV^{ω} that is strong enough to interpret the second-order theories of [19, 18, 49], which include a form of weak König's lemma?

The interpretation of Section 3 provides efficient translations between second-order and higher-order systems; and the Dialectica interpretation [6, 15, 27, 45] can be used to interpret the higher type theories in their quantifier-free counterparts. But the interpretation does not work at the first-order level. By internalizing cut-elimination arguments, it seems that one should be able to interpret $NPRA$ and $NERA$ efficiently in PRA and ERA. Is there an efficient means of interpreting NPV in PV? Or can one find specific counterexamples to show that this is not the case?

Is there a better way to treat equality in the theories presented here?

Are there interesting nonstandard versions of Feferman's theories of explicit mathematics?

There are more general questions, having to do with the formalization of mathematics in theories like the ones presented here. For example, what is required to formalize various parts of analysis? See [14, 17, 26, 27, 29, 30, 35, 37, 39, 40] for various approaches to answering this question. Do nonstandard theories provide a useful approach?

Can nonstandard theories like the ones presented here provide a perspicuous means of extracting polynomial bounds from proofs of theorems in analysis, as done by Kohlenbach [26, 27, 28, 29, 30]?

context $\forall^{st} u\ (\widehat{t} \not< u)$, expressing that t denotes a nonstandard element. We can more frugally take conditions to be represented pairs of the form $\langle \alpha, f \rangle$, where α is a predicate on N (represented by its characteristic function) and f is a function from N to N. Define such a pair $\langle \alpha, f \rangle$ to be a condition if it satisfies

$$\forall z\ \exists \omega\ (\alpha(\omega) \wedge f(\omega) \geq z)$$

expressing that the predicate α holds of values of ω making f arbitrarily large. The relation $\langle \beta, g \rangle \preceq \langle \alpha, f \rangle$ is defined by

$$\langle \beta, g \rangle \preceq \langle \alpha, f \rangle \equiv \forall \omega\ (\beta(\omega) \rightarrow \alpha(\omega) \wedge g(\omega) \leq f(\omega)),$$

and if $\langle \alpha, f \rangle$ and $\langle \beta, g \rangle$ are compatible conditions, then their greatest lower bound is given by $\langle \alpha \wedge \beta, \min(f, g) \rangle$. The forcing clauses for atomic formulae are now as follows:

- $p \Vdash t_1 = t_2 \equiv \exists z\ \forall \omega\ (\alpha(\omega) \wedge f(\omega) \geq z \rightarrow \widehat{t_1} = \widehat{t_2})$
- $p \Vdash t_1 < t_2 \equiv \exists z\ \forall \omega\ (\alpha(\omega) \wedge f(\omega) \geq z \rightarrow \widehat{t_1} < \widehat{t_2})$
- $p \Vdash st(t) \equiv \exists z\ \forall \omega\ (\alpha(\omega) \wedge f(\omega) \geq z \rightarrow \widehat{t} < z)$

We then have $\langle \top, \widehat{t} \rangle \Vdash \neg st(t)$; in particular, if id is the identity function, $\langle \top, id \rangle$ forces that ω is nonstandard. The reader can verify that the proof of Theorem 2.2 still goes through, mutatis mutandis.

Finally, in the case of $NPRA^{\omega}$, the translation can be simplified even further, provided one allows Σ_1 induction in the target theory. Conditions can be taken to be unary predicates p on N that hold for infinitely many values of ω,

$$\forall z\ \exists \omega \geq z\ p(\omega),$$

corresponding to sets in the Fréchet filter. These are the forcing clauses for atomic formulae:

- $p \Vdash t_1 = t_2 \equiv \exists z\ \forall \omega \geq z\ (p(\omega) \rightarrow \widehat{t_1} = \widehat{t_2})$
- $p \Vdash t_1 < t_2 \equiv \exists z\ \forall \omega \geq z\ (p(\omega) \rightarrow \widehat{t_1} < \widehat{t_2})$
- $p \Vdash st(t) \equiv \exists z\ \forall \omega \geq z\ (p(\omega) \rightarrow \widehat{t} < z)$

In other words, p forces $t_1 = t_2$ if and only if $\widehat{t_1} = \widehat{t_2}$ holds for all but finitely many values of ω satisfying p. Although the necessity of having Σ_1 induction in the interpreting theory weakens the result, it is perhaps surprising that such a straightfoward translation can be used to interpret nonstandard reasoning.

Σ_1 induction is required to verify that $\neg\neg st(t) \rightarrow st(t)$ is forced, as follows.

LEMMA B.1. *Over PRA^{ω}, Σ_1 induction is equivalent to the following principle*:

$$\exists z\ \forall y\ (f(y) \leq z) \rightarrow \exists x\ \forall y\ (f(y) \leq f(x)).$$

This principle expresses the fact that every bounded function on N has a least upper bound, and attains it. Only the forward direction of the lemma is required below, though the equivalence seems interesting in its own right.

PROOF. The contrapositive of the principle is equivalent to

$$\forall x\ \exists y\ (f(y) > f(x)) \rightarrow \forall z\ \exists y\ (f(y) > z). \tag{4}$$

We will show that Σ_1 induction is equivalent to (4), arguing in PRA^ω. For the forward direction, suppose $\forall x\ \exists y\ (f(y) > f(x))$. By induction on z, it is easy to show $\exists y\ (f(y) > z)$.

Conversely, suppose $\varphi(u, v)$ is a Δ_0 formula satisfying the two hypotheses of Σ_1 induction, $\exists v\ \varphi(0, v)$ and $\forall u\ (\exists v\ \varphi(u, v) \rightarrow \exists v\ \varphi(u + 1, v))$. We need to show $\forall u\ \exists v\ \varphi(u, v)$. Define $f(x)$ to be the greatest $w \leq x$ such that $\forall u < w\ \exists v \leq x\ \varphi(u, v)$. It is not hard to show that for every x, $\exists y\ (f(y) > f(x))$; if $f(x)$ is 0, this follows from the first hypothesis, and otherwise it follows from the second. By (4), for every u there is a y such that $f(y) > u$; the definition of f implies $\exists v\ \varphi(u, v)$. ⊣

LEMMA B.2. *Let t be any term. With the modified definition of forcing, $PRA^\omega + \Sigma_1$ induction proves the following: let p be any condition and let q be the predicate defined by*

$$q(\omega) \equiv p(\omega) \wedge \forall u < \omega\ (p(u) \rightarrow \widehat{t}(u) < \widehat{t}(\omega)).$$

Then if q is a condition, $q \Vdash \neg st(t)$.

PROOF. The idea is that q corresponds to a subset of p on which $\widehat{t}$ is strictly increasing as a function of ω. Suppose q is a condition, and let r be any predicate satisfying $r \preceq q$. It suffices to show that if $r \Vdash st(t)$, then r is not a condition.

So, suppose $r \Vdash st(t)$, i.e.

$$\exists z\ \forall \omega \geq z\ (r(\omega) \rightarrow \widehat{t}(\omega) < z). \tag{5}$$

Since $r \preceq q$ we know that $\widehat{t}$ is increasing on r, that is,

$$\forall u, v\ (r(u) \wedge r(v) \wedge u < v \rightarrow \widehat{t}(u) < \widehat{t}(v)). \tag{6}$$

Define f by

$$f(v) = \max_{u \leq v \wedge r(u)} \widehat{t}(u).$$

By (5) we have that f is bounded by z. Using the principle of Lemma B.1, there is a value u such that $\forall v\ (f(v) \leq f(u))$. But then (6) implies

$$\forall \omega > u\ \neg r(\omega),$$

so r is not a condition. ⊣

LEMMA B.3. *$PRA^\omega + \Sigma_1$ induction proves that $\neg\neg st(t) \rightarrow st(t)$ is forced.*

PROOF. Suppose $p \Vdash \neg\neg st(t)$. Then $\forall q \preceq p\ (q \not\Vdash \neg st(t))$. Define q as in the statement of Lemma B.2. Then $q \preceq p$, and if q is a condition then

$q \Vdash \neg st(t)$; so q is not a condition. This means we have $\exists z\ \forall \omega \geq z\ \neg q(\omega)$, i.e. for some z we have

$$\forall \omega \geq z\ (p(\omega) \to \exists u < \omega\ (p(u) \wedge \hat{t}(\omega) \leq \hat{t}(u))).$$

Since p is a condition, we can pick an $\omega \geq z$ satisfying $p(\omega)$, and let $v = \max_{u \leq \omega \wedge p(u)} \hat{t}(u)$. Then we have

$$\forall \omega \geq z\ (p(\omega) \to t(\omega) \leq v)$$

which implies $p \Vdash st(t)$. ⊣

With the modified forcing definition, it is easy to show that $\neg st(\omega)$ is forced. So, in the end, we can conclude that whenever $NPRA^{\omega}$ proves a formula φ, $PRA^{\omega} + \Sigma_1$ induction proves that φ is forced.

REFERENCES

[1] Miklos Ajtai, *The complexity of the pigeonhole principle*, ***Proceedings of the IEEE 29th annual symposium on foundations of computer science***, 1988, pp. 346–355.

[2] Jeremy Avigad, *Formalizing forcing arguments in subsystems of second-order arithmetic*, ***Annals of Pure and Applied Logic***, vol. 82 (1996), no. 2, pp. 165–191.

[3] ———, *Interpreting classical theories in constructive ones*, ***The Journal of Symbolic Logic***, vol. 65 (2000), no. 4, pp. 1785–1812.

[4] ———, *Ordinal analysis without proofs*, ***Reflections on the foundations of mathematics (Stanford, California, 1998)***, Lecture Notes in Logic, vol. 15, Association for Symbolic Logic, Urbana, Illinois, 2002, pp. 1–36.

[5] ———, *Saturated models of universal theories*, ***Annals of Pure and Applied Logic***, vol. 118 (2002), pp. 219–234.

[6] Jeremy Avigad and Solomon Feferman, *Gödel's functional ("Dialectica") interpretation*, In Buss [13], pp. 337–405.

[7] Jeremy Avigad and Jeffrey Helzner, *Transfer principles in nonstandard intuitionistic arithmetic*, ***Archive for Mathematical Logic***, vol. 41 (2002), no. 6, pp. 581–602.

[8] Jon Barwise (editor), ***The handbook of mathematical logic***, North-Holland, Amsterdam, 1977.

[9] Michael J. Beeson, ***Foundations of constructive mathematics***, Springer, Berlin, 1985.

[10] Samuel R. Buss, ***Bounded arithmetic***, Bibliopolis, Naples, 1986, A reprinting of the author's 1985 Princeton University dissertation.

[11] ———, *A conservation result concerning bounded theories and the collection axiom*, ***Proceedings of the American Mathematical Society***, vol. 100 (1987), pp. 709–716.

[12] ———, *First-order proof theory of arithmetic*, In Buss [13].

[13] Samuel R. Buss (editor), ***The handbook of proof theory***, North-Holland, Amsterdam, 1998.

[14] Rolando Chuaqui and Patrick Suppes, *Free-variable axiomatic foundations of infinitesimal analysis: a fragment with finitary consistency proof*, ***The Journal of Symbolic Logic***, vol. 60 (1995), no. 1, pp. 122–159.

[15] Stephen A. Cook and Alasdair Urquhart, *Functional interpretations of feasibly constructive arithmetic*, ***Annals of Pure and Applied Logic***, vol. 63 (1993), pp. 103–200.

[16] Solomon Feferman, *Theories of finite type related to mathematical practice*, In Barwise [8], pp. 913–971.

[17] António M. Fernandes and Fernando Ferreira, *Groundwork for weak analysis*, ***The Journal of Symbolic Logic***, vol. 67 (2002), no. 2, pp. 557–578.

[18] ———, *Basic applications of weak König's lemma in feasible analysis*, ***Reverse mathematics 2001*** (S. Simpson, editor), Lecture Notes in Logic, vol. 22, AK Peters, 2005, this volume, pp. 175–188.

[19] Fernando Ferreira, *A feasible theory for analysis*, ***The Journal of Symbolic Logic***, vol. 59 (1994), pp. 1001–1011.

[20] ———, *A note on a result of Buss concerning bounded theories and the collection scheme*, ***Portugaliae Mathematica***, vol. 52 (1995), no. 3, pp. 331–336.

[21] Kurt Gödel, ***Collected works***, vol. II, Oxford University Press, New York, 1990, Solomon Feferman et al., editors.

[22] Robert Goldblatt, ***Lectures on the hyperreals: an introduction to nonstandard analysis***, Springer, New York, 1998.

[23] Petr Hájek and Pavel Pudlák, ***Metamathematics of first-order arithmetic***, Springer, Berlin, 1993.

[24] Richard Kaye, ***Models of Peano arithmetic***, Clarendon, Oxford, 1991.

[25] Ker-I Ko, ***Complexity theory of real functions***, Birkhäuser, Boston, 1991.

[26] Ulrich W. Kohlenbach, *Effective moduli from ineffective uniqueness proofs: an unwinding of de La Vallée Poussin's proof for Chebycheff approximation*, ***Annals of Pure and Applied Logic***, vol. 64 (1993), pp. 27–94.

[27] ———, *Analyzing proofs in analysis*, ***Logic: From foundations to applications: European logic colloquium '93*** (W. Hodges et al., editors), Clarendon Press, Oxford, 1996, pp. 225–260.

[28] ———, *Mathematically strong subsystems of analysis with low rate of growth of provably recursive functionals*, ***Archive for Mathematical Logic***, vol. 36 (1996), pp. 31–71.

[29] ———, *Proof theory and computational analysis*, ***Third workshop on computation and approximation (Comprox III, Birmingham, 1997)***, Elsevier, Amsterdam, 1998, 34 pp. (electronic).

[30] ———, *Foundational and mathematical uses of higher types*, ***Reflections on the foundations of mathematics (Stanford, California, 1998)***, Lecture Notes in Logic, vol. 15, Association for Symbolic Logic, Urbana, Illinois, 2002, pp. 92–116.

[31] Georg Kreisel, *Axiomatizations of nonstandard analysis that are conservative extensions of classical systems of analysis*, ***Applications of model theory to algebra, analysis, and probability*** (W. A. J. Luxemburg, editor), Holt, Rinehart, and Winston, 1967, pp. 93–106.

[32] Ieke Moerdijk and Erik Palmgren, *Minimal models of Heyting arithmetic*, ***The Journal of Symbolic Logic***, vol. 62 (1997), pp. 1448–1460.

[33] Edward Nelson, *Internal set theory: a new approach to nonstandard analysis*, ***Bulletin of the American Mathematical Society***, vol. 83 (1977), pp. 1165–1198.

[34] ———, ***Radically elementary probability theory***, Princeton University, 1987.

[35] Erik Palmgren, *Developments in constructive nonstandard analysis*, ***The Bulletin of Symbolic Logic***, vol. 4 (1998), pp. 233–272.

[36] ———, *An effective conservation result for nonstandard analysis*, ***Mathematical Logic Quarterly***, vol. 46 (2000), pp. 17–23.

[37] Stephen G. Simpson, ***Subsystems of second order arithmetic***, Springer, Berlin, 1998.

[38] Stephen G. Simpson and Rick L. Smith, *Factorization of polynomials and Σ_1^0 induction*, ***Annals of Pure and Applied Logic***, vol. 31 (1986), pp. 289–306.

[39] Richard Sommer and Patrick Suppes, *Finite models of elementary recursive nonstandard analysis*, ***Notas de la Sociedad de Matematica de Chile***, vol. 15 (1996), no. 1, pp. 73–95.

[40] ———, *Dispensing with the continuum*, ***Journal of Mathematical Psychology***, vol. 41 (1997), no. 1, pp. 3–10.

[41] Thomas Strahm, *Theories with self-application and computational complexity*, ***Information and Computation***, vol. 185 (2003), pp. 263–297.

[42] Kazuyuki Tanaka, *Non-standard analysis in WKL_0*, ***Mathematical Logic Quarterly***, vol. 43 (1997), pp. 401–412.

[43] A.S. Troelstra, ***Metamathematical investigation of intuitionistic arithmetic and analysis***, Lecture Notes in Mathematics, vol. 344, Springer, Berlin, 1973.

[44] ———, *Aspects of constructive mathematics*, In Barwise [8], pp. 973–1052.

[45] ———, *Introductory note to 1958 and 1972*, In Gödel [21], pp. 217–241.

[46] ———, *Realizability*, In Buss [13], pp. 407–473.

[47] Alex Wilkie, *Modèles non standard de l'arithmétique, et complexité algorithmique*, ***Modèles non standard en arithmétique et théorie des ensembles***, Publications Mathémathiques de l'Université Paris VII, vol. 22, 1985.

[48] Alan R. Woods, *Approximating the structures accepted by a constant depth circuit or satisfying a sentence—a nonstandard approach*, ***Logic and random structures (New Brunswick, New Jersey, 1995)***, American Mathematical Society, Providence, Rhode Island, 1997, pp. 109–130.

[49] Takeshi Yamazaki, *Reverse mathematics and weak systems of* 0-1 *strings for feasible analysis*, ***Reverse mathematics 2001*** (S. Simpson, editor), Lecture Notes in Logic, vol. 22, AK Peters, 2005, this volume, pp. 394–401.

DEPARTMENT OF PHILOSOPHY
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PA 15213, USA
E-mail: avigad@cmu.edu
URL: http://www.andrew.cmu.edu/~avigad

follows from the statement of the theorem, together with the more usual proof of the theorem from the same set of axioms provides the precise knowledge that these axioms are in some sense necessary to prove a theorem of ordinary mathematics. In such a case we have a very complete answer to the Main Question.

In this paper we consider the axiomatic strength required to prove the equivalence of three common notions of compactness for a complete separable metric space:

i) total boundedness;
ii) Heine-Borel compactness;
iii) sequential compactness.

In section two we present the definition of a complete separable metric space and some relevant results and examples. The third section considers the basic topological definitions of open and closed sets. This section also presents the precise definitions of the three common notions of compactness and considers the axiomatic strength required to prove their equivalence. In the fourth and final section we present the Stone-Weierstrass theorem for compact complete separable metric spaces.

§**2. Metric spaces.** Within RCA_0 we define a (code for a) *complete separable metric space* to consist of a set $A \subset \mathbb{N}$ together with a function $d : A \times A \to \mathbb{R}$ such that for all $a, b, c \in A$:

i) $d(a, a) = 0$;
ii) $0 \leqslant d(a, b) = d(b, a)$;
iii) $d(a, c) \leqslant d(a, b) + d(b, c)$.

Now let (A, d) be a code for a complete separable metric space, as above. We define, again within RCA_0, a *point in the completion* $\hat{A}$ to be a function $f : \mathbb{N} \to A$ such that

$$\forall n\, \forall i\ [d(f(n), f(n+i)) < 2^{-n}].$$

The idea here is that (A, d) is a code for the complete separable metric space $\hat{A}$ consisting of all such points. For example, $\mathbb{R} = \widehat{\mathbb{Q}}$ under the usual pseudometric. Of course $\hat{A}$ does not formally exist within RCA_0. A point $f : \mathbb{N} \to A$ will be denoted by $x = \langle a_n : n \in \mathbb{N}\rangle$ where $a_n = f(n)$ and we will sometimes use the notation $(x)_n$ for a_n. Two points $x = \langle a_n : n \in \mathbb{N}\rangle$ and $y = \langle b_n : n \in \mathbb{N}\rangle$ are said to be *equal* if $\forall n\ [d(a_n, b_n) < 2^{-n+1}]$. A *sequence* of points of $\hat{A}$ is a function $f : \mathbb{N} \to \hat{A}$ and is denoted by $\langle x_n : n \in \mathbb{N}\rangle$ where $x_n = f(n)$. We extend the pseudometric d on A to a pseudometric $\hat{d}$ on $\hat{A}$ by defining

$$\hat{d}\big(\langle a_n : n \in \mathbb{N}\rangle, \langle b_n : n \in \mathbb{N}\rangle\big) = \langle c_{n,n} : n \in \mathbb{N}\rangle,$$

where

$$\langle c_{n,k} : k \in \mathbb{N}\rangle = d(a_{n+3}, b_{n+3}).$$

Where no confusion will result, d will be used to denote both d and $\hat{d}$. We embed A into $\hat{A}$ by identifying the element $a \in A$ with the point $x_a \in \hat{A}$ defined by $x_a = \langle a : n \in \mathbb{N} \rangle$. Thus, under this embedding, A is a countable dense subset of $\hat{A}$. We also note that $\hat{A}$ is complete, in the following sense:

LEMMA 2.1 (RCA_0). *Given a sequence* $\langle x_n : n \in \mathbb{N} \rangle$ *of points in* $\hat{A}$ *such that* $\forall n \, \forall i \, [d(x_n, x_{n+i}) < 2^{-n}]$, *there exists a point* $x \in \hat{A}$ *such that* $x = \lim_{n\to\infty} x_n$, *i.e.* $\lim_{n\to\infty} d(x, x_n) = 0$.

PROOF. For $n \in \mathbb{N}$, let $x_n = \langle a_{n,k} : k \in \mathbb{N} \rangle$, where each $a_{n,k} \in A$. Let $x = \langle a_{n+3,n+3} : n \in \mathbb{N} \rangle$. It is easy to show that $x \in \hat{A}$ and that $\lim_{n\to\infty} d(x, x_n) = 0$. ⊣

We will also need the following technical result:

LEMMA 2.2 (RCA_0). *The following are equivalent*:

i) ACA_0;
ii) *If* $f : \mathbb{N} \to \mathbb{N}$ *is total and one-to-one, then* $\exists X \, \forall n \, [n \in X \leftrightarrow \exists m \, (f(m) = n)]$.

PROOF. See [11]. ⊣

Given a sequence of complete separable metric spaces, we can define their product as follows: let $\langle A_n : n \in \mathbb{N} \rangle$ be a sequence of codes for complete separable metric spaces $\hat{A}_n$. For each $n \in \mathbb{N}$, let c_n be the least element (in the usual ordering of $\mathbb{N}$) of A_n. Let P be the set of all finite sequences $\langle a_0, \ldots, a_i \rangle$ with $a_k \in A_k$, $0 \leqslant k \leqslant i$. For sequences $\langle a_0, \ldots, a_i \rangle$ and $\langle b_0, \ldots, b_j \rangle$ in P, define

$$d(\langle a_0, \ldots, a_i \rangle, \langle b_0, \ldots, b_j \rangle) = \sum_{n=0}^{k} \frac{1}{2^{n+1}} \cdot \frac{d(a_n, b_n)}{1 + d(a_n, b_n)},$$

where $k = \max(i, j)$, $a_n = c_n$ for $n > i$ and $b_n = c_n$ for $n > j$. Then P codes the complete separable metric space $\prod_{\mathbb{N}} \hat{A}_n$ with

$$d(x, y) = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \cdot \frac{d(x_n, y_n)}{1 + d(x_n, y_n)}$$

for $x = \langle x_n : n \in \mathbb{N} \rangle$ and $y = \langle y_n : n \in \mathbb{N} \rangle$ with $x_n, y_n \in \hat{A}_n$. Thus, within RCA_0, we have the *Baire Space* $\mathbb{N}^{\mathbb{N}} = \prod_{\mathbb{N}} \mathbb{N}$.

§3. The topology of complete separable metric spaces. Within RCA_0 we define a code for an *open ball* in $\hat{A}$, a complete separable metric space with code (A, d), to be an ordered pair (x, ε) where $x \in \hat{A}$ and $\varepsilon \in \mathbb{R}^+$ (the positive reals). We say that a point $y \in \hat{A}$ is an *element of the ball* (x, ε) if $d(x, y) < \varepsilon$, in which case we will write $y \in (x, \varepsilon)$. An open ball is a *basic open set* if it is of the form (a, r) where $a \in A$ and $r \in \mathbb{Q}^+$ (the positive rationals). A code for an *open set* U is a (possibly empty) sequence of basic open sets $\langle (a_n, r_n) : n \in \mathbb{N} \rangle$.

We say that a point $x \in \hat{A}$ *belongs to* U if there is a basic open set $(a,r) \in U$ such that x is an element of (a,r), in which case we will write $x \in U$. We have the following basic results about open sets:

LEMMA 3.1 (RCA_0). *Let $\langle U_n : n \in \mathbb{N}\rangle$ be a sequence of open sets in a complete separable metric space $\hat{A}$. Then $U = \bigcup_{\mathbb{N}} U_n$ is an open set.*

LEMMA 3.2 (RCA_0). *Let $\langle U_i : 0 \leqslant i \leqslant n\rangle$ be a finite sequence of open sets in a complete separable metric space $\hat{A}$. Then $\bigcap_0^n U_i$ is open.*

The proofs may be found in [4]. For open balls (x,ε) and (y,δ) we will write $(x,\varepsilon) < (y,\delta)$ to mean $\varepsilon < \delta - d(x,y)$ and $(x,\varepsilon) \leqslant (y,\delta)$ to mean $\varepsilon \leqslant \delta - d(x,y)$. Thus $(x,\varepsilon) \leqslant (y,\delta)$ implies that the open ball (x,ε) is included in the open ball (y,δ) while $(x,\varepsilon) < (y,\delta)$ allows that this inclusion may be proper.

There are two natural definitions of a closed set. The distinctions between them form the basis of [4]. The first natural definition of a closed set is that it is the complement of an open set. Thus we define a (code for a) *closed set* C to be a (possibly empty) sequence of basic open sets $\langle (a_n, r_n) : n \in \mathbb{N}\rangle$, and say that a point x in a complete separable metric space $\hat{A}$ *belongs to* C if $d(a_n, x) \geqslant r_n$ for all $n \in \mathbb{N}$. Note that a code for a closed set may also be regarded as a code for the open set which is its complement. We then easily obtain:

LEMMA 3.3 (RCA_0). *Let $\langle C_n : n \in \mathbb{N}\rangle$ be a sequence of closed sets in a complete separable metric space $\hat{A}$. Then $C = \bigcap_{\mathbb{N}} C_n$ is also a closed set.*

LEMMA 3.4 (RCA_0). *Let $\langle C_n : 0 \leqslant i \leqslant n\rangle$ be a finite sequence of closed subsets of $\hat{A}$. Then $C = \bigcup_0^n C_i$ is closed.*

Since every closed subset of a complete separable metric space is itself a complete separable metric space, a second natural definition of a closed set is that it is the closure of a countable set of points. We therefore define a (code for a) *separably closed set* $\bar{S}$ to consist of a sequence $S = \langle x_n : n \in \mathbb{N}\rangle$ of points from a complete separable metric space $\hat{A}$. We say a point $x \in \hat{A}$ *belongs to* $\bar{S}$ if $\forall r \in \mathbb{Q}^+ \ \exists n\ [d(x, x_n) < r]$. We will occasionally write $\overline{\{x_n : n \in N\}}$ for $\bar{S}$. Note that the definition of a separably closed set is equivalent to that of a closed subspace of $\hat{A}$ given in [5].

Two notions related to these definitions of closed set are investigated by Giusto and Simpson in [9]. Specifically, in RCA_0 we define a closed or separably closed subset C of a complete separable metric space $\hat{A}$ to be *located* if there exists (a code for) the continuous *distance function*

$$f(x) = d(x, C) = \inf\{d(x,y) \mid y \in C\}.$$

We say that C is *weakly located* if the predicate

$$\exists \varepsilon > 0\ \left[B(s, r+\varepsilon) \cap C = \emptyset\right]$$

is Σ_1^0.

In the context of weak subsystems of second order arithmetic there is an important distinction between the two definitions of closed set: relatively strong axioms are required to prove their equivalence. Specifically we have the following:

THEOREM 3.5 (RCA_0). *The following are equivalent*:

i) ACA_0;

ii) *If* $\bar{S}$ *is a separably closed subset of a complete separable metric space then* $\bar{S}$ *is a closed set*.

PROOF. See [4]. ⊣

THEOREM 3.6 (RCA_0). *The following are equivalent*:

i) Π^1_1-CA_0;

ii) *If* C *is a closed subset of a complete separable metric space, then* C *is a separably closed set*.

PROOF. See [4]. A corrected proof due to Jeffrey L. Hirst of the statement that any closed subset of a complete separable metric space is separably closed implies ACA_0 is soon to appear. ⊣

Thus we see that for an arbitrary complete separable metric space, the equivalence of the two notions of closed set requires, and is equivalent to, Π^1_1-CA_0. In the setting of compact spaces (as defined below) this equivalence can be proved in the weaker system ACA_0, as we note below. We also note that the analogues of Lemmas 3.3 and 3.4 hold, in Π^1_1-CA_0, for separably closed sets. We may also define a *separably open set* to be the complement of a separably closed set, for details see [4].

Let $\hat{A}$ be a complete separable metric space. We say that $\hat{A}$ is *compact* if there is a sequence $\langle\langle x_{n,i} : i \leqslant i_n\rangle : n \in \mathbb{N}\rangle$ of finite sequences of points in $\hat{A}$ such that, for all $x \in \hat{A}$ and $n \in \mathbb{N}$, there is an $i \leqslant i_n$ with $d(x, x_{n,i}) < 2^{-n}$. Thus $\hat{A}$ is compact if it is (effectively) totally bounded. We say that $\hat{A}$ is *Heine-Borel compact* if, given any open cover $\langle U_n : n \in \mathbb{N}\rangle$ of $\hat{A}$, there is a $k \in \mathbb{N}$ such that $\langle U_n : n \leqslant k\rangle$ also covers $\hat{A}$. Finally, we say that $\hat{A}$ is *sequentially compact* if, given any sequence $\langle x_n : n \in \mathbb{N}\rangle$ of points in $\hat{A}$, there is a subsequence $\langle x_{n_k} : k \in \mathbb{N}\rangle$, $n_0 < n_1 < \cdots < n_k < \cdots$, such that $\lim_{k\to\infty} x_{n_k}$ exists. One reason for choosing to define a totally bounded space to be compact is that we can, for example, prove within RCA_0 that the unit interval is (in this sense) compact. For the other two notions of compactness we have the following:

LEMMA 3.7 (RCA_0). *The following are equivalent*:

i) WKL_0;

ii) $[0, 1]$ *is Heine-Borel compact*.

PROOF. See [11]. ⊣

LEMMA 3.8 (RCA_0). *The following are equivalent*:

i) ACA_0;

ii) $[0, 1]$ *is sequentially compact*.

PROOF. See [11]. ⊣

In contrast to Theorem 3.6 above, in a compact complete separable metric space we have the following result:

THEOREM 3.9 (RCA_0). *The following are equivalent*:

i) ACA_0;
ii) *If $\hat{A}$ is compact then every closed set in $\hat{A}$ is separably closed*;
iii) *If $\hat{A}$ is compact then every separably closed set in $\hat{A}$ is closed.*

PROOF. See [4]. ⊣

If we also consider the notions of located or weakly located, we have the following set of results:

THEOREM 3.10 (RCA_0). *The following are equivalent*:

i) ACA_0;
ii) *If $\hat{A}$ is compact then every closed set in $\hat{A}$ is located*;
iii) *If $\hat{A}$ is compact then every separably closed set in $\hat{A}$ is located*;
iv) *If $\hat{A}$ is compact then every separably closed set in $\hat{A}$ is weakly located*;
v) *If $\hat{A}$ is compact then every closed and weakly located set in $\hat{A}$ is located*;
vi) *If $\hat{A}$ is compact then every closed and weakly located set in $\hat{A}$ is separably closed.*

PROOF. See [9]. ⊣

THEOREM 3.11 (RCA_0). *If $\hat{A}$ is compact then every closed and located set in $\hat{A}$ is separably closed.*

PROOF. See [9]. ⊣

THEOREM 3.12 (RCA_0). *The following are equivalent*:

i) WKL_0;
ii) *If $\hat{A}$ is compact then every closed and separably closed set in $\hat{A}$ is located*;
iii) *If $\hat{A}$ is compact then every closed and separably closed set in $\hat{A}$ is weakly located*;
iv) *If $\hat{A}$ is compact then every closed set in $\hat{A}$ is weakly located.*

PROOF. See [9]. ⊣

For more details on the notions of located and weakly located sets, see [9].

In a standard presentation of real analysis one shows that the three notions of compactness defined above are, in fact, equivalent. With limited axiomatic strength this is not the case until we reach ACA_0, as we demonstrate in the remainder of this section.

THEOREM 3.13 (RCA_0). *The following are equivalent*:

i) ACA_0;
ii) *Every Heine-Borel compact space $\hat{A}$ is compact.*

$n \leqslant m$ there is an $i_n \leqslant k_n$ such that $d(x, x_{n,i_n}) < 2^{-n}$. Let $\sigma \in \mathbb{N}^{<\mathbb{N}}$ be such that $\mathrm{lh}(\sigma) = m + 1$ and $\sigma(n) = i_n$ for each $n \leqslant m$. Note that for all $i, j < \mathrm{lh}(\sigma)$,

$$d\left(x_{i,\sigma(i)}, x_{j,\sigma(j)}\right) \leqslant d\left(x_{i,\sigma(i)}, x\right) + d\left(x, x_{j,\sigma(j)}\right) < 2^{-i} + 2^{-j}$$

so that for all $k \in \mathbb{N}$, and in particular all $k < \mathrm{lh}(\sigma)$, $(d(x_{i,\sigma(i)}, x_{j,\sigma(j)}))_k < 2^{-i} + 2^{-j} + 2^{-k}$. Therefore, since $\sigma \notin T$, there must exist $i, j, k < \mathrm{lh}(\sigma)$ such that $(d(x_{i,\sigma(i)}, a_j))_k < r_j - 2^{-i} - 2^{-k}$ and hence $(d(x_{i,\sigma(i)}, a_j))_k < r_j - 2^{-i}$. Then

$$d\left(x, a_j\right) \leqslant d\left(x, x_{i,\sigma(i)}\right) + d\left(x_{i,\sigma(i)}, a_j\right) < r_j,$$

i.e., $x \in (a_j, r_j)$ and $j \leqslant m$. Thus (i) implies (ii), as desired.

To see that (ii) implies (i), we note that [0, 1] is compact (take the sequence $\langle\langle k2^{-n} : k \leqslant 2^{-n}\rangle : n \in \mathbb{N}\rangle$) and then apply Lemma 3.7. ⊣

LEMMA 3.16 (RCA_0). *If $\hat{A}$ is sequentially compact then it is Heine-Borel compact.*

PROOF. Let $U = \langle(b_n, r_n) : n \in \mathbb{N}\rangle$ be an open cover of $\hat{A}$. If U has no finite subcover, it follows that for each $k \in \mathbb{N}$ there is some $x \in \hat{A}$ such that $x \in \sim \bigcup_{n \leqslant k} (b_n, r_n)$, i.e. $d(x, b_n) \geqslant r_n$ for all $n \leqslant k$. Let $x = \langle a_i : i \in \mathbb{N}\rangle$. Then for each $n \leqslant k$,

$$\begin{aligned} r_n &\leqslant d(x, b_n) \\ &\leqslant d(x, a_k) + d(a_k, b_n) \\ &< 2^{-k} + d(a_k, b_n). \end{aligned}$$

It follows that $d(a_k, b_n) > r_n - 2^{-k}$ for all $n \leqslant k$ and hence that $\{a \in A \mid \forall n \leqslant k\ (d(a, b_n) > r_n - 2^{-k})\} \neq \emptyset$. Let

$$c_k = \mu a \in A\ (\forall n \leqslant k\ (d(a, b_n) > r_n - 2^{-k}))$$

and consider the sequence $\langle c_k : k \in \mathbb{N}\rangle$. If $\hat{A}$ is sequentially compact, there is a subsequence $\langle c_{k_i} : i \in \mathbb{N}\rangle$, $k_i < k_{i+1}$ which converges to some $y \in \hat{A}$. Fix $n \in \mathbb{N}$ and $\varepsilon > 0$ and let i_0 be such that for all $i \geqslant i_0$, $d(y, c_{k_i}) < \varepsilon$. Then for all $k \geqslant \max(n, k_{i_0})$,

$$\begin{aligned} r_n - 2^{-k} &< d(b_n, c_k) \\ &\leqslant d(b_n, y) + d(y, c_k) \\ &< d(b_n, y) + \varepsilon. \end{aligned}$$

Thus $d(b_n, y) > r_n - 2^{-k} - \varepsilon$ for all $k \geqslant \max(n, k_{i_0})$ so that $d(b_n, y) \geqslant r_n - \varepsilon$. Since $\varepsilon > 0$ was arbitrary, it follows that $d(b_n, y) \geqslant r_n$, i.e. $y \notin (b_n, r_n)$. As n was arbitrary, we have $y \notin \bigcup_{\mathbb{N}} (b_n, r_n) = U$, contradicting the assumption that U is an open cover of $\hat{A}$. Thus if U is an open cover of $\hat{A}$, it must have a finite subcover, as desired. ⊣

THEOREM 3.17 (RCA_0). *The following are equivalent*

i) ACA_0;

ii) *Every sequentially compact space $\hat{A}$ is compact.*

PROOF. For (i) implies (ii), assume ACA_0 and suppose further that $\hat{A}$ is sequentially compact. By Lemma 3.16, $\hat{A}$ is Heine-Borel compact and therefore, by Theorem 3.13, $\hat{A}$ is compact.

For (ii) implies (i), let $f : \mathbb{N} \to \mathbb{N}$ be 1-1 and let $\hat{S} = \{2^{-f(k)} : k \in \mathbb{N}\} \cup \{0\}$, as in Theorem 3.13. Clearly $\hat{S}$ is sequentially compact so that, if (ii) holds, ACA_0 follows as in Theorem 3.13. ⊣

We now establish a result for sequential compactness corresponding to that of Theorem 3.15 for Heine-Borel compactness:

THEOREM 3.18 (RCA_0). *The following are equivalent.*

i) ACA_0;

ii) *If $\hat{A}$ is compact then it is sequentially compact.*

PROOF. For (i) implies (ii), suppose that $\hat{A}$ is compact as witnessed by the sequence $\langle\langle x_{n,i} : i \leqslant i_n\rangle : n \in \mathbb{N}\rangle$. Let $\langle y_k : k \in \mathbb{N}\rangle$ be a sequence of points in $\hat{A}$ and define a set $T \subseteq \mathbb{N}^{<\mathbb{N}}$ by:

$$\begin{aligned}\sigma \in T \text{ iff } &\forall j < \mathrm{lh}(\sigma)\,\big[\sigma(j) < i_j\big]\\ &\wedge \forall i, j < \mathrm{lh}(\sigma)\,\big[d\left(x_{i,\sigma(i)}, x_{j,\sigma(j)}\right) < 2^{-i} + 2^{-j}\\ &\qquad \wedge \forall m\, \exists k \geqslant m\, \left(d\left(y_k, x_{i,\sigma(i)}\right) < 2^{-i}\right)\big].\end{aligned}$$

Note that T exists within ACA_0. Intuitively, $\sigma \in T$ iff σ describes a finite sequence of overlapping balls of decreasing radius, each of which contains infinitely many points of the sequence $\langle y_k : k \in \mathbb{N}\rangle$. Clearly T is a bounded tree and, using the pigeonhole principle, it is easy to see that T is infinite. Therefore, by Lemma 3.14, there is a path $f \in \mathbb{N}^{\mathbb{N}}$ through T. Consider the sequence $\langle x_{n+1,f(n+1)} : n \in \mathbb{N}\rangle$. As in the proof of Theorem 3.15, there is an $x \in \hat{A}$ such that $x = \lim_{n\to\infty} x_{n+1,f(n+1)}$. Without loss of generality we may assume that $d(x, x_{n+1,f(n+1)}) < 2^{-(n+1)}$. Define a sequence $\langle k_n : n \in \mathbb{N}\rangle$ by recursion, as follows:

$$\begin{aligned}k_0 &= \mu k\left[d\left(y_k, x_{1,f(1)}\right) < 2^{-1}\right]\\ k_{n+1} &= \mu k\left[k > k_n \wedge d\left(y_k, x_{n+1,f(n+1)}\right) < 2^{-(n+1)}\right].\end{aligned}$$

Note that this sequence exists as f is a path through T and so the ball $(x_{n+1,f(n+1)}, 2^{-(n+1)})$ contains infinitely many points of the sequence $\langle y_k : k \in \mathbb{N}\rangle$. Then for all $n \in \mathbb{N}$ we have:

$$\begin{aligned}d(x, x_{n_k}) &\leqslant d\left(x, x_{n+1,f(n+1)}\right) + d\left(x_{n+1,f(n+1)}, y_{k_n}\right)\\ &< 2^{-(n+1)} + 2^{-(n+1)} = 2^{-n}.\end{aligned}$$