Notas de la Sociedad Matematica de Chile, 15, 1996, pp. 73-95.

FINITE MODELS OF ELEMENTARY RECURSIVE NONSTANDARD ANALYSIS

RICHARD SOMMER AND PATRICK SUPPES

ABSTRACT. This paper provides a new proof of the consistency of the formal system presented by Chuaqui and Suppes in [2, 9]. First, a simpler, yet stronger, system, called Elementary Recursive Nonstandard Analysis, ERNA, will be provided. Indeed, it will be shown that ERNA proves all of the theorems of the Chuaqui and Suppes system. Then a finitary consistency proof of ERNA will be given; in particular, we will show that *PRA*, the system of primitive recursive arithmetic, which is generally recognized as capturing Hilbert's notion of finitary, proves the consistency of ERNA. From the consistency proof we can extract a constructive method for obtaining finite approximations of models of nonstandard analysis. We present an isomorphism theorem for models that are finite substructures of infinite models.

1. INTRODUCTION

This paper continues and extends the development of a constructive system of nonstandard analysis begun by Chuaqui and Suppes in [2, 9]. The approach is meant to provide a foundation that is close to the mathematical practice characteristic of theoretical physics. For detailed elaboration of this viewpoint see [9]. Perhaps the most important point to mention here is that many standard theorems are weakened in our constructive framework to proofs of approximate equality rather than exact equality. But an infinitesimal difference is as good as equality for physical purposes.

Date: January 3, 1996.

This paper is dedicated to the memory of Rolando Chuaqui. The joint work with Rolando and the second author of the present paper goes back more than a decade, and our friendship many more years than that. In about 1985, we began working on formal rules of derivation for a computer-based calculus course. A long technical report of this period was reduced to manageable size in Chuaqui and Suppes [1]. The equational deductive system of this 1990 paper was first presented at an International Conference on Computer Logic in Tallinn, Estonia, in 1988. Out of this concrete focus on computer implementation we came to the deeper problem of developing a system of constructive nonstandard analysis, which was first presented at the IX Latin American Symposium on Mathematical Logic in Bahia Blanca, Argentina in 1991, and published as [9]. At the time of Rolando's death in the spring of 1994, we were working (by e-mail) on the first notes for the isomorphism theorem presented in this paper, and we were planning to meet at Stanford in June, 1994, following a visit of mine to Santiago in January, 1994. His sudden and unexpected death left a still unfilled void for many of us, both in Chile and in the United States. His wonderfully energetic and optimistic spirit is much missed.

This paper will not deal with the development of mathematics in weak systems of nonstandard analysis. The reader is referred to [4] for a general development of nonstandard analysis, and to [2, 9] to see how a significant amount of the analysis that is relevant to applications in the physical sciences, can be developed in ERNA. Although the work here is a continuation of the work presented in [2, 9], this paper is self-contained.

In part, the intention of this paper is to present a system of nonstandard analysis called *elementary recursive nonstandard analysis*, ERNA, which is a simplified, and more versatile, version of the system presented by Chuaqui and Suppes in [2, 9]. As in [2], we give a finitary consistency proof for our system; the proof here is free of many of the technical details present in the proof given in [2], hopefully making this proof more transparent. Also, we will clarify, precisely, to what extent this consistency proof is finitistic; in the terminology of proof theory, we measure the proof theoretic strength of ERNA.

The claim, stated above, that ERNA is more versatile than the system of Chuaqui and Suppes follows from the fact that ERNA allows for a form of *definition by recursion*. Functions, such as factorial and summation, that are defined using recursion, are built into the system in [2, 9] individually; there is also an implicit use of definition by recursion in the development of calculus given in that paper. Incorporating definition by recursion directly into the system allows the simplification of not having to include axioms for several individual functions, and it justifies further use of this method for defining additional functions.

Another feature of ERNA that differs from the system of Chuaqui and Suppes is that ERNA contains a constant symbol $\uparrow$ that is intended to denote an "undefined element." Since the system includes division, we need to deal with the possibility that 0 will occur in the denominator. Also, some of the other function symbols of the language denote functions that are not total. The constant symbol $\uparrow$ denotes the value of a function for the cases that the function is "undefined"; for example, $1/0 = \uparrow$.

In order to have a finitary consistency proof for the system it is necessary to restrict the form of definition by recursion; it is well known that arithmetic with full primitive recursion cannot be finitisticly proven consistent. In ERNA recursion is restricted in the same way as in the development of Kalmar elementary arithmetic; hence the terminology *elementary* recursive nonstandard analysis (cf. [6]). This will be explained further in Section 2 where the details of ERNA are presented. In Section 3, we show that, indeed, ERNA contains the system presented in [2].

The consistency proof for ERNA is given in Section 4. The proof will use Herbrand's theorem, and it applies a construction modeled after techniques introduced by Paris and Kirby in [5], and applied by the first author in [7] and [8]. What will be noted in Section 4, and then elaborated on in later sections, is the fact that the proof entails an algorithm for constructing a kind of finite model for the system, which is also the case for the consistency proof in [2]. A discussion of the limitations of this kind of construction is presented in Section 5.

In Section 6, we go on to prove a theorem that asserts that the substructure determined by the interpretation of a finite set of terms in a "reasonably sound" infinite model of nonstandard analysis is isomorphic to finite substructures of the standard rational numbers. In that section we will discuss how this yields a partial isomorphism from the *physically continuous* to the *physically discrete*. The existence of such an isomorphism is surprising, but it supports the strong physical intuition that no experiments can distinguish between any physical quantities, even space and time, being continuous or discrete at a fine enough level. Philosophically, we can say that the continuum may be real for Platonists, but it can nowhere be unequivocally identified in the real world of physical experiments.

Although the consistency proof is very constructive, the proof of the isomorphism theorem is not. This will be discussed in Section 7. But, in addition to negative results regarding the inability to construct the finite models referred to in the isomorphism theorem, we present a positive result, showing that in a certain special case such a model can be constructed. In particular, we present a very simple example taken from a physical problem.

2. Elementary Recursive Nonstandard Analysis

2.1. The language of ERNA. ERNA is formulated in a language with

1. **Variables:** $v_0, v_1, \ldots$.
2. **Relation symbols:** Inf, $\mathcal{N}$, $=$, and $\leq$. (Inf denotes the set of infinitesimals and $\mathcal{N}$ denotes the set of natural numbers.)
3. **Individual constant symbols:** 0, 1, ν_0, ϵ_0, and $\uparrow$. (ν_0 denotes an infinite natural number and ϵ_0 denotes its reciprocal; $\uparrow$ is used to denote the value of an "undefined term" (such as $1/0$), and $x = \uparrow$ is read "x is undefined".)
4. **Function symbols:**
   a. $|\ |$, $\lceil\ \rceil$, $+$, $-$, $\cdot$, $\div$, and exp . ($|\ |$ denotes absolute value, $\lceil\ \rceil$ denotes the least integer greater than or equal to, and exp denotes exponentiation with natural number powers.)
   b. An l-ary function symbol $\pi_{l,i}$ for each pair of positive integers l and i, where $i \leq l$. ($\pi_{l,i}$ denotes a projection function.)
   c. An m-ary function symbol $\min_\varphi$ for each quantifier-free formula φ with $m+1$ distinct free variables, that does not contain terms involving min. ($\min_\varphi$ denotes a *minimum operator* for the least natural number satisfying φ. The restriction that min cannot occur in φ is necessary for the system to have a finitary consistency proof; this will be explained further in Section 2.4.)
   d. An $m+1$-ary function symbol $\mathrm{rec}^b_{\sigma\tau}$ for each positive integer b and each pair σ, τ of terms, with arities m and $m+2$, respectively, that do not involve min. *Note:* σ and τ may contain occurrences of rec.

Also, the arity of a term is the number of distinct free variables in that term. ($\mathrm{rec}^b_{\sigma\tau}$ denotes the function obtained from σ and τ by *definition by recursion*. The parameter, b, is used to bound the growth rate of the function. This will be explained in more detail shortly. The restriction that min cannot occur in the terms σ and τ comes about for the same reason as the similar restriction in c, above.)

2.2. Conventions and Definitions.

1. Throughout this paper we will use vector notation, such as $\vec{x}$, to denote sequences of variables, terms, etc. For example, $\vec{x}$ is understood to abbreviate $x_1, \ldots, x_l$, for some l that can be determined from the context.
2. $x \exp y$ will be abbreviated by x^y.
3. A formula is **internal** if it does not contain occurrences of Inf. An **external** formula may or may not contain occurrences of Inf.
4. The variables i, j, k, l, m, and n, will range, exclusively, over $\mathcal{N}$; for example, $\varphi(n)$ is an abbreviation for

$$\mathcal{N}(v_i) \implies \varphi(v_i)$$

for the appropriate variable, v_i.

5. $x \approx y$ will be used to denote $\mathrm{Inf}(x - y)$, and will be read "x is approximately equal to y". Thus $x \approx 0$ will be used to denote $\mathrm{Inf}(x)$, but may still be read "x is infinitesimal."
6. $x \approx \infty$ is used to denote $x \neq 0 \wedge \mathrm{Inf}(1/x)$ and is read x **is infinite.** Also, $x \not\approx \infty$ is used to denote $x = 0 \vee \neg\mathrm{Inf}(1/x)$, and is read x **is finite.**
7. For τ a term, $(\tau \downarrow)$ will be used to denote $\tau \neq \uparrow$, and will be read "τ is defined". Also, $(\tau \uparrow)$ may be written in place of $\tau = \uparrow$, and is read "τ is undefined."

2.3. The axioms of ERNA.

(N) **Natural number** axioms:

1. $\mathcal{N}(0)$ and if $\mathcal{N}(x)$ then $\mathcal{N}(x+1)$.
2. If $\mathcal{N}(x)$ then either $x = 0$ or $\mathcal{N}(x-1)$.
3. If $\mathcal{N}(x)$ then $x \geq 0$.
4. $\mathcal{N}(\nu_0)$.

(I) **Infinity** axioms:

1. The sum of two infinitesimals is an infinitesimal:

$$x \approx 0 \wedge y \approx 0 \rightarrow x + y \approx 0.$$

2. The product of an infinitesimal and a finite number is an infinitesimal:

$$x \approx 0 \wedge y \not\approx \infty \rightarrow xy \approx 0.$$

3. Infinitesimals are finite:

$$x \approx 0 \rightarrow x \not\approx \infty.$$

4. If $|x|$ is less than an infinitesimal then x is an infinitesimal:

$$|x| \leq y \wedge y \approx 0 \rightarrow x \approx 0.$$

5. The sum of two finite numbers is finite:

$$x \not\approx \infty \wedge y \not\approx \infty \rightarrow x + y \not\approx \infty.$$

6. ϵ_0 is an infinitesimal:

$$\epsilon_0 \approx 0.$$

7. $\epsilon_0 = 1/\nu_0$.

(F) Axioms asserting that the elements, other than $\uparrow$, constitute an ordered field of characteristic 0 with absolute value function.

(A) Archimedean axiom: If $x \neq \uparrow$ then $\mathcal{N}(\lceil x \rceil)$ and $\lceil x \rceil - 1 < x \leq \lceil x \rceil$. (This is called the *Archimedean axiom* since it implies the Archimedean principle that asserts that for each x there is a natural number n such that $x \leq n$.)

(E) Recursive defining equations for **exponentiation** for natural number powers: if $x \neq \uparrow$ then

$$x^0 = 1 \text{ and } x^{n+1} = x^n \cdot x.$$

(P) Axioms for the **projection functions**: if, for $j = 1, \ldots, l$, $x_l \neq \uparrow$ then

$$\pi_{l,i}(\vec{x}) = x_i.$$

The next group of axioms, (R), formulates the version of definition by recursion that is allowed in ERNA. To make this definition more transparent, we provide some motivation. First consider the usual equations for definition by recursion; the following define, by recursion, a function f, from functions g and h.

$$(*) \quad f(0, \vec{x}) = g(\vec{x}) \text{ and } f(n+1, \vec{x}) = h(f(n, \vec{x}), n, \vec{x}).$$

Although these equations are easier to read if the parameters $\vec{x}$ are suppressed, we exhibit the parameters since it will be important to keep track of them in the version of recursion allowed in ERNA. Although recursion in the above form is very useful, and simple to present, it cannot be justified finitisticly (in the sense that if it is allowed as an axiom, the resulting system will not have a finitary consistency proof). This can be remedied by putting bounds on the growth rates of the functions obtained using this definition. In particular, if we require functions given by $(*)$ to be bounded in growth rate by, say, an exponential function, then the system has a finitary consistency proof (as will be demonstrated in Section 4).

Observing that exponentiation is built into ERNA, in axiom (E), it is clear that ERNA allows for finite iterations of the exponential function. In particular, the following function is given by a term in ERNA,

$$x \mapsto 2_k^x =_{\text{def}} \underbrace{2^{2^{2^{\cdots 2^x}}}}_{k\ 2's},$$

where k is a natural number. In fact, it is straightforward to show that terms that do not involve min or rec, are bounded by a function of the form $x \mapsto 2_k^x$ for some fixed k.

By using terms of the form 2_k^x as bounds on the functions defined by recursion, we can insure that all terms, not involving min, are bounded by such a term. This is a key point to carry out the consistency proof. Also it should be noted that this is a suitable bound for the functions, defined by recursion, that come up in practice.

Since division, and hence taking reciprocals, is built into the language of ERNA, we need to also have the growth rate bounds be given in terms of reciprocals of the arguments of the function that is being defined. In particular, we want all terms, not involving min, to satisfy the property,

$$\text{if } |\vec{x}| \in [1/a, a] \cup \{0\} \text{ then } |\tau(\vec{x})| \in [1/b, b] \cup \{0, \uparrow\},$$

where $b = 2_k^a$, for a fixed k only depending on τ.

Also, since we only have exponentiation for natural number powers, we need to use the function $\lceil\ \rceil$ in the specification of the growth bounds. These considerations complicate the statement of definition by recursion given below in axiom (R), but the reader should be able to readily see that (R) is just a version of (*) that incorporates the growth bounds discussed above. The term σ appears in place of g, τ appears in place of h, and $\text{rec}_{\sigma\tau}^b$ appears in place of f. To simplify notation in the statement of our bounds, we write $\max |\vec{x}|, |\vec{x}|^{-1}, n$ to abbreviate,

$$\max\{\lceil |x_j| \rceil : j = 1, \dots, l\} \cup \{\lceil \frac{1}{|x_j|} \rceil : x_j \neq 0\} \cup \{n\}.$$

(R) Axioms for definition by recursion:

1. If $1/2_b^{\max |\vec{x}|, |\vec{x}|^{-1}, 0} \leq |\sigma(\vec{x})| \leq 2_b^{\max |\vec{x}|, |\vec{x}|^{-1}, 0}$, or $\sigma(\vec{x}) = \uparrow$, then
$$\text{rec}_{\sigma\tau}^b(0, \vec{x}) = \sigma(\vec{x})$$
otherwise $\text{rec}_{\sigma\tau}^b(0, \vec{x}) = 0$. The assignment of this "default value" insures, first, that the definition is correct when $\sigma(\vec{x}) = 0$, and second, that the necessary growth rate bounds are maintained.
2. If
$$1/2_b^{\max |\vec{x}|, |\vec{x}|^{-1}, n+1} \leq |\tau(\text{rec}_{\sigma\tau}^b(n, \vec{x}), n, \vec{x})| \leq 2_b^{\max |\vec{x}|, |\vec{x}|^{-1}, n+1},$$
or if $\tau(\text{rec}_{\sigma\tau}^b(n, \vec{x}), n, \vec{x}) = \uparrow$, then
$$\text{rec}_{\sigma\tau}^b(n+1, \vec{x}) = \tau(\text{rec}_{\sigma\tau}^b(n, \vec{x}), n, \vec{x}),$$
otherwise $\text{rec}_{\sigma\tau}^b(n+1, \vec{x}) = 0$.

(IM) Quantifier-free internal minimum: For each quantifier-free internal formula φ, that does not involve min, if $\varphi(n,\vec{x})$ then

$$\mathcal{N}(\min_\varphi(\vec{x})),\ \min_\varphi(\vec{x}) \le n, \text{ and } \varphi(\min_\varphi(\vec{x}),\vec{x}).$$

and if $\neg\varphi(\min_\varphi(\vec{x}),\vec{x})$ then $\min_\varphi(\vec{x}) = 0$. This axiom specifies that $\min_\varphi$ is given by:

$$\min_\varphi(\vec{x}) = \begin{cases} \text{the least } n \text{ such that } \varphi(n,\vec{x}) & \text{if } \exists n \varphi(n,\vec{x}), \\ 0 & \text{otherwise.} \end{cases}$$

(EM) Quantifier-free external induction: This is the same as (IM) except for two modifications. First, this axiom refers to external, rather than intermal, formulas, and second, in this axiom n is restricted to be finite; i.e., in this axiom $\min_\varphi(\vec{x})$ is the least $n \not\approx \infty$ such that $\varphi(n,\vec{x})$, if there is such an n, and is 0, otherwise, where φ is a quantifier free external formula not involving min.

(U) Undefined Terms:

1. $(0\downarrow)$, $(1\downarrow)$, $(\nu_0\downarrow)$, and $(\epsilon_0\downarrow)$
2. $(|x|\downarrow) \leftrightarrow (x\downarrow) \leftrightarrow (\lceil x\rceil\downarrow)$.
3. $(x+y\downarrow) \leftrightarrow (x-y\downarrow) \leftrightarrow (xy\downarrow) \leftrightarrow ((x\downarrow)\wedge(y\downarrow))$.
4. $(x \div y\downarrow) \leftrightarrow ((x\downarrow)\wedge(y\downarrow)\wedge y\neq 0)$.
5. $(x^y\downarrow) \leftrightarrow ((x\downarrow)\wedge(y\downarrow)\wedge\mathcal{N}(y))$.
6. $(\pi_{l,i}(\vec{x})\downarrow) \leftrightarrow ((x_1\downarrow)\wedge\ldots\wedge(x_l\downarrow))$.
7. $\neg\mathcal{N}(x) \implies (\text{rec}^b_{\sigma\tau}(x,\vec{y})\uparrow)$.
8. $(\min_\varphi(\vec{x})\downarrow) \leftrightarrow ((x_1\downarrow)\wedge\ldots\wedge(x_l\downarrow))$.

Note that (U7) is of a different form than the other (U) axioms. This is justified by the fact that rec will use the other function symbols of the language when defining a function; so, as a result of (R), the functions defined using recursion will become "undefined" exactly when they should.

2.4. Remarks about ERNA.

1. Note that (P), (R), (IM), and (EM) are axiom schemata; i.e., they consist of infinitely many axioms.
2. Important for our consistency proof is that all of the axioms of ERNA are quantifier-free, with free variables. This is clear from an inspection of the statement of these axioms as given Section 2.3, except for the field axioms, which are not listed. The field axioms that assert the existence of additive and multiplicative inverses are typically given as existential statements. In ERNA, we suppose they are given by:
   - $x+(-x)=0$, where $-x =_{\text{def}} (0-x)$, and
   - If $x\neq 0$, then $x\cdot x^{-1}=1$, where $x^{-1} =_{\text{def}} 1/x$.

   Also, since we have the constants 0 and 1, we do not need existentially quantified axioms for the existence of additive and multiplicative identity elements.
3. As is the case in formalizations of number theory, having a minimum operator is equivalent to having induction; in particular, axioms (IM)

and (EM) are equivalent to induction axioms for quantifier-free, internal and external, respectively, formulas that do not involve min. In the case of (EM), this is an induction axiom for the "finite" natural numbers.

4. It is worth mentioning that in many formalizations of fragments of number theory, definition by recursion can be proved from induction. This is not the case here since we do not have a coding apparatus for coding sequences of elements of the universe that would allow us to carry out the usual proof from number theory.
5. As was mentioned in Section 2.1, when introducing the symbols $\min_\varphi$, the restriction of not allowing min to appear in φ is necessary to have a finitary consistency proof. The operator resulting from the removal of this restriction is equivalent to a minimum operator for formulas with natural number quantifiers (note: $\exists n\varphi(n, \vec{x})$ is equivalent to $\varphi(\min_\varphi(\vec{x}), \vec{x})$.) The strength of the resulting system would be the same as Peano Arithmetic, which does not have a finitary consistency proof.

3. The Suppes and Chuaqui system is contained in ERNA

The system in [2] has the following function symbols that are not contained in ERNA:

$$\delta, \text{li}, \max_\tau, \sum, \text{ and } !.$$

The system in [2] also has symbols, and axioms, for differentiation and integration, but it is shown (early in [2, Section 3]) that the system is a conservative extension of the system without differentiation and integration, and thus we need only concern ourselves with the latter system.

Mainly, to show that the system in [2] is contained in ERNA, we need to show that the above functions are definable, and their defining axioms are provable in ERNA. In the process of doing so we will exhibit a list of basic functions that are definable in ERNA. By definable we mean that there is a term of the language that (provably in ERNA) has the properties of the function.

1. The identity function *id* is definable in ERNA. In particular, *id* is simply $\pi_{1,1}$.
2. For each closed term τ there are constant functions $C_{k,\tau}$ of each arity k with value τ. In particular, $\pi_{k+1,k+1}(x_1, \ldots, x_k, \tau)$ is such a function.
3. There is a definable function ζ such that,

$$\zeta(x) = \begin{cases} 1 & \text{if } x = 0, \text{ and} \\ x & \text{otherwise.} \end{cases}$$

In particular, ζ is obtained from the function r, defined by recursion,

$$r(n) =_{\text{def}} \begin{cases} 0 & \text{if } n = 0, \text{ and} \\ 1 & \text{otherwise.} \end{cases}$$

(r is $\mathrm{rec}^1_{0C_{2,1}}$), by taking,

$$\zeta(x) =_{\mathrm{def}} (x+1) - r(\lceil |x| \rceil).$$

Note: it will be useful to use ζ in the event a denominator is potentially 0; an example of this occurs in the next item in this list.

4. The function δ_1 given by

$$\delta_1(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

is definable by

$$\delta_1(x) = \frac{|x| + x}{2\zeta(x)}.$$

5. The function, $d_{\rho\sigma\tau}$, *defined by cases* from the terms ρ, σ and τ, is given by

$$d_{\rho\sigma\tau}(\vec{x}) = \begin{cases} \sigma(\vec{x}) & \text{if } \rho(\vec{x}) \geq 0 \\ \tau(\vec{x}) & \text{otherwise,} \end{cases}$$

is definable in ERNA. In particular,

$$d(\vec{x}) = \delta_1(\rho(\vec{x}))\sigma(\vec{x}) + (1 - \delta_1(\rho(\vec{x})))\tau(\vec{x}).$$

Now, on to the functions of the system of [2]:

6. The function δ, from [2], given by

$$\delta(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{otherwise,} \end{cases}$$

is definable using definition by cases, so is definable in ERNA.

7. $\max_\tau$, from [2], is given by

$$\max_\tau(n) = \begin{cases} 1 & \text{if } n = 1. \\ \begin{cases} n+1 & \text{if } \tau(\max_\tau(n)) < \tau(n+1) \\ \max_\tau(n) & \text{if } \tau(\max_\tau(n)) \geq \tau(n+1). \end{cases} & \text{if } n > 1. \end{cases}$$

Since this is a combination of definition by recursion and definition by cases, it is justified in ERNA.

8. ! is defined by recursion: $1! = 1$ and $(n+1)! = n! \cdot (n+1)$.
9. $\sum$ is defined by recursion:

$$\sum_{k=1}^{1} \tau = \tau(1) \text{ and } \sum_{k=1}^{n+1} \tau = \left(\sum_{k=1}^{n} \tau \right) + \tau(n+1).$$

10. $\mathrm{li}(x)$, an integer greater than x, can be taken to be $\lceil x \rceil$.

Hence, all of the functions of the system of Chuaqui and Suppes in [2], that are not directly built into ERNA, are definable in ERNA, so ERNA contains that system.

4. The consistency of ERNA

This proof will use Herbrand's theorem in the same way as the consistency proof presented in [2]. In particular, an algorithm will be given for assigning rational number values val(τ) to each term τ in a fixed finite set of terms $\mathcal{T}$; we extend the rationals by including an element to represent an "undefined value," i.e., the interpretation of $\uparrow$. The val-assignment will then be used to assign truth values to atomic formulas. This will be done in such a way that any axiom of ERNA, that only involves terms in $\mathcal{T}$, will be assigned *true*. Furthermore, all of the logical equality axioms will be true in this assignment. From Herbrand's theorem, we conclude ERNA is consistent. Also, it will be shown that the val-assignment, in fact the entire consistency proof, can be carried out in primitive recursive arithmetic (PRA), and hence, ERNA has a finitary consistency proof.

It is useful to view the assignment of the val function as the construction of a finite model. In particular, the elements of the range of val, when restricted to our fixed finite set of terms, make up the universe of what can be construed as a finite approximation of a model of the theory. This will be explained and discussed in much greater detail in the following sections (it is the focus of the later sections of this paper); at this point, we mention this only to describe how we are using this terminology to explain and motivate the assignment of val.

Now we describe the form of the finite sets of terms that will be handled in our construction. Suppose Φ is a finite set of quantifier free formulas that do not contain terms involving $\min_\varphi$. Let $\mathcal{T}_i^\Phi$ denote the set of terms with depth less than or equal to i that only involve $\min_\varphi$ if φ is in Φ, where the **depth** $d(\tau)$ of a term τ is defined as follows

- If τ is an individual constant then $d(\tau) = 0$.
- $d(\lceil\tau\rceil) = d(|\tau|) = d(\tau) + 1$.
- $d(\sigma \star \tau) = \max\{d(\sigma), d(\tau)\} + 1$, if $\star$ is $+$, $-$, $\cdot$, $\div$, or exp .
- $d(\pi_{l,i}(\vec{\tau})) = l + \max\{d(\tau_1), \ldots, d(\tau_l)\}$.
- $d(\mathrm{rec}^b_{\sigma\tau}(\vec{\rho})) = b + d(\sigma) + d(\tau) + \max\{d(\rho_1), \ldots, d(\rho_l)\} + 1$.
- $d(\min_\varphi(\vec{\tau})) = \max\{d(\tau_i) : i = 1, \ldots, l\} + 1$.

Note that $\varphi(k, \vec{\tau})$ may have occurrences of terms of depth strictly greater than $d(\min_\varphi(\vec{\tau}))$. Clearly, $\mathcal{T}_i^\Phi$ is finite; in fact if t_i denotes the cardinality of $\mathcal{T}_i^\Phi$, and if A is one greater than the sum of the arities of the (finitely many) functions that occur in terms in $\mathcal{T}_i^\Phi$, then

$$t_i \leq 5^{A^i}.$$

This follows from the fact that there are five constants in the language, so five terms of depth 0, and the fact

$$t_{j+1} \leq t_j^A.$$

We suppose a maximum depth D and a finite set of formulas Φ is fixed. The val-assignment will be given to terms in $\mathcal{T}_D^\Phi$ inductively on the depth of terms, and it will satisfy the following *homomorphism* properties:

Val1 For all functions, f, of ERNA, *except* $\min_\varphi$, ϵ_0, *and* ν_0:

$$\mathrm{val}(f(\tau_1,\ldots,\tau_l)) = f(\mathrm{val}(\tau_1),\ldots,\mathrm{val}(\tau_l)).$$

Individual constants are treated as 0-ary functions.

Val2 For all relations R of ERNA *except* Inf:

$$\mathrm{val}(R(\tau_1,\ldots,\tau_l)) = true \iff R(\mathrm{val}(\tau_1),\ldots,\mathrm{val}(\tau_l)).$$

These properties determine the val assignment in almost all cases, and since the axioms (N1–3), (F), (A), (E), (P), (R), and (U) all hold in the rationals, together with $\dagger$, **Val1** and **Val2** insure that our truth assignment will assign *true* to all instances of these axioms that only involve terms in $\mathcal{T}_D^\Phi$. For example, suppose 0, τ, and $\tau+1$ are all in $\mathcal{T}_D^\Phi$, and consider the following instance of (N1)

$$\mathcal{N}(0) \wedge (\mathcal{N}(\tau) \implies \mathcal{N}(\tau+1)).$$

By **Val1**, $\mathrm{val}(0) = 0$, and, by **Val2**, noting that $\mathcal{N}(0)$ is true, $\mathrm{val}(\mathcal{N}(0)) =$ *true*, so the first conjunct of the axiom is assigned true. If $\mathrm{val}(\tau)$ is not a natural number then, by **Val2**, $\mathrm{val}(\mathcal{N}(\tau)) =$ *false*, and so the second conjunct is also true. If $\mathrm{val}(\tau)$ is a natural number then, using **Val1**, $\mathrm{val}(\tau+1) = \mathrm{val}(\tau)+1$ is also a natural number, and, by **Val2**, the second conjunct is again true.

It remains to show how to assign val to ν_0, ϵ_0, $\min_\varphi$, and Inf, so that (N4), (I), (IM), and (EM) hold. More specifically, we need to deal with the following

a. $\mathrm{val}(\nu_0)$ needs to be assigned a finite natural number value so that (N4) and (I6) are satisfied. To satisfy (I7), we will take $\mathrm{val}(\epsilon_0) = 1/\mathrm{val}(\nu_0)$.

b. Values have to be assigned to terms involving $\min_\varphi$ *in a constructive way*. It is important to note that we cannot assign val to terms involving $\min_\varphi$ by using **Val1** as we did with the other functions. This is because the computation of $\min_\varphi$ is not constructive — it involves an unbounded search. The val-assignment for $\min_\varphi$ will be made constructive by bounding the search; although $\mathrm{val}(\min_\varphi(\vec{\tau}))$ may not receive its true value (in the sense of the standard rationals), it will receive a value that will lead to an assignment of *true* to the relevant instances of (IM) and (EM).

c. Truth values need to be assigned to Inf so that the axioms (I) hold. Also, we will need these assignments to assign values to the external minimum operator.

Satisfying a–c, above, involves selecting sequences of standard natural numbers, a_i, b_i, c_i, and d_i, for $i = 0,\ldots,D$. The construction carried out here is modeled after one that first appeared in [5], and was developed further in [7] and [8]. How these sequences are used in the val-assignment, and how they are defined, will be explained in detail below. Since we are simultaneously working to satisfy several axioms, and, at the same time, we are working to keep our construction finitistic, there are many considerations that need to be taken into account in this construction. We will proceed in the following steps:

Step 1: The val-assignment for ϵ_0, ν_0, Inf, and $\min_\varphi$ will be given.
Step 2: Assuming certain properties of these sequences, the axioms of ERNA are shown to hold.
Step 3: The construction of the sequences will be described.
Step 4: The construction will be shown to be finitistic.

In the following paragraph we give an informal description of some of the intuitions behind the construction. This description is intended to give the reader a better picture of the construction; it is not intended to constitute an essential part of the proof.

We view the val-assignment and the construction of the sequences of a's, b's, c's, and d's as a D stage construction, where in stage i terms of depth i are handled assuming that terms of depth less than i were assigned values at earlier stages. The numbers a_i, b_i, c_i, and d_i are associated with stage i of the construction. All of the a's are less than all of the b's which are less than all of the c's which are less than all of the d's, and the a_i's and the c_i's are increasing sequences, and the b_i's and d_i's are decreasing sequences (this is expressed in P1, below in step 2). The interval $[b_i, c_i]$ represents the "infinite" part of the model at stage i in the sense that if τ is of depth less than or equal to i and $b_i \leq |\text{val}(\tau)| \leq c_i$, then the formula $\tau \approx \infty$ is assigned *true*. Implicit in this assertion is the fact that c_i serves as an upper bound for $|\text{val}(\tau)|$, when τ has depth i. So $\{0\} \cup [\frac{1}{c_i}, \frac{1}{b_i}]$ represents the infinitesimal part of the model at stage i in the sense that if $\tau \in \mathcal{T}_i^\Phi$, and $|\text{val}(\tau)| \in \{0\} \cup [\frac{1}{c_i}, \frac{1}{b_i}]$ then $\text{val}(\tau \approx 0) = \textit{true}$. And the interval $[\frac{1}{a_i}, a_i]$ represents the finite, non-infinitesimal part of the model; i.e., if $\tau \in \mathcal{T}_i^\Phi$ and $|\text{val}(\tau)| \in [\frac{1}{a_i}, a_i]$ then $\text{val}(\tau \approx 0) = \textit{false}$ and $\text{val}(\tau \approx \infty) = \textit{false}$. In most cases, Val1 is used to assign values to terms. P2, below, insures that there will be "enough space" to carry this out; this is explained in step 3.

Step 1: Now the remaining cases of the val-assignment are described.

Val3 $\text{val}(\nu_0) = b_0$ and $\text{val}(\epsilon_0) = 1/b_0$.
Val4 If $\tau \in \mathcal{T}_i^\Phi$ then $\text{val}(\text{Inf}(\tau)) = \textit{true}$ if and only if $|\text{val}(\tau)| \leq \frac{1}{b_i}$.

To increase readability for the val-assignment for terms involving $\min_\varphi$, we introduce two abbreviations. Let $\text{val}(\vec{\tau})$ denote the sequence $\text{val}(\tau_1), \ldots, \text{val}(\tau_l)$, and let

$$(\mu n \leq m)\varphi =_{\text{def}} \begin{cases} \text{the least } n \leq m \text{ such that } \varphi(n) & \text{if } (\exists n \leq m)\varphi, \\ 0 & \text{otherwise.} \end{cases}$$

Val5 If $\min_\varphi(\vec{\tau})$ has depth i, and φ is internal then we set

$$\text{val}(\min_\varphi(\vec{\tau})) = (\mu n \leq d_i)\varphi(n, \text{val}(\vec{\tau}))$$

Val6 If $\min_\varphi(\vec{\tau})$ has depth i, and φ is external then we set

$$\mathrm{val}(\min_\varphi(\vec{\tau})) = (\mu n \le b_i)\varphi(n, \mathrm{val}(\vec{\tau})).$$

Step 2: Here we show that the axioms of ERNA, associated with $\min_\varphi$, ϵ_0, ν_0, and Inf, are satisfied by our val-assignment. For that we will assert some properties of the sequences a_i, b_i, c_i, and d_i. For now the reader will have to take on faith that sequences with these properties can be constructed; we handle that in step 3.

P1 $0 < a_{i-1} \le a_i \le b_i \le b_{i-1} \le c_{i-1} \le c_i \le d_i \le d_{i-1}$.
P2 If $\tau \in \mathcal{T}_i^\Phi$ then $\frac{1}{|\mathrm{val}(\tau)|}, |\mathrm{val}(\tau)| \in [0, a_i] \cup [b_i, c_i] \cup \{\uparrow\}$.
P3 If $0 < i \le D$ then $2a_{i-1} < a_i$.
P4 If $0 < i \le D$ then $(b_i)^2 < b_{i-1}$.

(N1–3), (F), (A), (E), (P), (R), and (U). As was pointed out in the discussion immediately following the statements of **Val1** and **Val2**, these axioms are all an immediate consequence of **Val1** and **Val2**.

In the discussion of the remaining axioms, we assume that, for all terms τ, the rules given in (U) can not be used to show $\mathrm{val}(\tau) = \uparrow$; i.e., we assume that all terms are "defined." The undefined cases are handled in a straightforward way.

(N4) Since b_0 is a natural number, **Val2** and **Val3** insure that $\mathcal{N}(\nu_0)$ is assigned *true*.

(I1) Suppose val(Inf(τ)) = *true*, val(Inf(σ)) = *true*, and $\tau + \sigma \in \mathcal{T}_D^\Phi$. Then $\tau, \sigma \in \mathcal{T}_{D-1}^\Phi$, and, using **Val4**, $|\mathrm{val}(\tau)|, |\mathrm{val}(\sigma)| \le 1/b_{D-1}$. Thus, using **Val1**, and properties of absolute value, $|\mathrm{val}(\tau+\sigma)| \le 2/b_{D-1} \le 1/b_D$; this last inequality follows from P4 with D in place of i and noting, from P1, P3, and P4, $b_D \ge 2$. Hence val(Inf($\tau + \sigma$)) = *true*.

(I2) Suppose val(Inf(τ)) = *true*, val(σ) = 0 or val(Inf($1/\sigma$)) = *false*, and $\sigma \cdot \tau \in \mathcal{T}_D^\Phi$. In the case val($\sigma$) = 0, **Val1** implies val($\sigma \cdot \tau$) = 0, so, by **Val4**, val(Inf($\sigma \cdot \tau$)) = *true*. Otherwise, note $\tau, \sigma \in \mathcal{T}_{D-1}^\Phi$. Also, val(Inf($1/\sigma$)) = *false* implies $|\mathrm{val}(\sigma)| \le b_{D-1}$. Using P2, we have $|\mathrm{val}(\sigma)| \le a_{D-1}$. Also, $|\mathrm{val}(\tau)| \le 1/b_{D-1}$, so, using **Val1** and properties of absolute value, $|\mathrm{val}(\sigma\cdot\tau)| \le a_{D-1}/b_{D-1} \le 1/b_D$; this last inequality follows from P4 noting also that, from P1, $a_{D-1} \le b_D$. Hence val(Inf($\sigma \cdot \tau$)) = *true*.

(I3) We need to show that if val(τ) $\ne 0$ then Inf(τ) and Inf($1/\tau$) cannot both be assigned true. Noting **Val4**, this is the same as saying $|\mathrm{val}(\tau)| \le 1/b_i$ and $|\mathrm{val}(1/\tau)| \le 1/b_i$ cannot both hold. This is clearly the case by **Val1** and the fact $b_i > 1$.

(I4) This readily follows from **Val4**.

(I5) Suppose val($\tau \not\approx \infty$) = *true*, val($\sigma \not\approx \infty$) = *true*, and $\tau + \sigma \in \mathcal{T}_D^\Phi$. Then $\tau, \sigma \in \mathcal{T}_{D-1}^\Phi$, and, using **Val4** and P2, $|\mathrm{val}(\tau)|, |\mathrm{val}(\sigma)| \le a_{D-1}$. Thus, using **Val1**, and properties of absolute value, $|\mathrm{val}(\tau+\sigma)| \le 2a_{D-1} \le a_D$; this last inequality follows from P3 with D in place of i. So val($\tau + \sigma \not\approx \infty$) = *true*.

(I6) This follows readily from **Val3** and **Val4**.

(I7) Is immediate from **Val3**.

(IM) Suppose $\min_\varphi(\vec{\tau})$ has depth $i \leq D$, and suppose $\varphi(\text{val}(\sigma), \text{val}(\vec{\tau}))$ holds for some $\sigma \in \mathcal{T}_D^\Phi$. Then since all val-assignments, for the elements of $\mathcal{T}_D^\Phi$, are less than or equal to c_D, and, by P1, $c_D \leq d_D \leq d_i$, we have $\varphi(n, \text{val}(\vec{\tau}))$ holds for some $n < d_i$, and hence, by **Val5**, $\text{val}(\min_\varphi(\vec{\tau})) =$ the least such n. This insures that (IM) holds. On the other hand, suppose $\varphi(\text{val}(\sigma), \text{val}(\vec{\tau}))$ fails for all $\sigma \in \mathcal{T}_D^\Phi$, then $\varphi(\text{val}(\min_\varphi(\vec{\tau})), \text{val}(\vec{\tau}))$ fails, so, by **Val5**, $\text{val}(\min_\varphi(\vec{\tau})) = 0$, and again (IM) is seen to hold.

(EM) The argument is to similar to that for (IM), but b_i is used in place of d_i.

Step 3: In order to define the a_i's, b_i's, c_i's and d_i's, we will use the following hierarchy of functions defined on natural numbers.

- $f_0(x) = 2_D^x$.
- $f_{n+1}(x) = f_n^T(x)$, where $T = 2t_D + 3$.

Note: the exponent, T, indicates T-fold iteration; i.e., $f_n^T(x) = \underbrace{f_n(\cdots f_n(f_n(x))\cdots)}_{T\ f_n's}$.

In addition to properties P1–5, we will show that the a_i's, b_i's, c_i's, and d_i's satisfy the following properties. These are needed in the induction hypothesis to continue the construction from stage to stage.

P5 For $i = 1, \ldots, D$, $b_i = f_{D-i}(a_i)$.
P6 For $i = 1, \ldots, D$, $d_i = f_{D-i}(c_i)$.

By setting $a_0 = 1$, $b_0 = f_D(a_0)$, $c_0 = b_0$, and $d_0 = f_D(c_0)$, it is straightforward to check that P1–7 hold for $i = 0$.

Suppose that $a_0, \ldots, a_i$, $b_0, \ldots, b_i$, $c_0, \ldots, c_i$, and $d_0, \ldots, d_i$ have been selected such that P1–7 hold, and suppose $i < D$. We need to show how to select a_{i+1}, b_{i+1}, c_{i+1}, and d_{i+1} so that P1-7 are satisfied with $i+1$ in place of i. Using P5, and the definition of the f_n's we see that the interval $[a_i, b_i]$ can be partitioned into $2t_D + 3$ intervals of the form

$$(*) \quad (f^j_{D-(i+1)}(a_i), f^{j+1}_{D-(i+1)}(a_i)]$$

for $j = 0, \ldots, T-1 = 2t_D + 2$. Similarly, using P6, the interval $[c_i, d_i]$ can likewise be partitioned. Now, consider a term $f(\vec{\tau}) \in \mathcal{T}_{i+1}^\Phi$. If f is any function of ERNA, except $\min_\varphi$, we can compute $\text{val}(f(\vec{\tau}))$. **Val5** and **Val6** cannot be used to compute $\text{val}(\min_\varphi(\vec{\tau}))$ since we do not yet know what d_{i+1} and b_{i+1} are. In order to find b_{i+1} and d_{i+1}, we compute

$$n_\varphi = (\mu n \leq d_i)\varphi(n, \text{val}(\vec{\tau})),$$

if φ is internal, and

$$n_\varphi = (\mu n \leq b_i)\varphi(n, \text{val}(\vec{\tau})),$$

if φ is external. Let S be the set of all of the numbers n_φ together with all of the values of $f(\vec{\tau})$ in the cases where f is not $\min_\varphi$. Also, close S under taking reciprocals; i.e., if $x \in S$ and $x \neq 0$, put $1/x \in S$. Since there are t_D terms in $\mathcal{T}_D^\Phi$, there are at most $2t_D$ elements in S. Now note that there are

$2t_D + 3$ intervals of the form given in (*); using the pigeon-hole principle, pick one that has empty intersection with S. Since there are at least three more intervals of the form (*) than there are elements of S, we can suppose $1 \leq j \leq 2t_D + 1$. For such a j, let

$$(**) \quad a_{i+1} = f^j_{D-(i+1)}(a_i) \text{ and } b_{i+1} = f^{j+1}_{D-(i+1)}(a_i).$$

The numbers c_{i+1} and d_{i+1} are chosen in a similar manner. Now we show that P1–7 hold with i replaced by $i+1$.

From the construction, $[a_{i+1}, b_{i+1}]$ is a subinterval of $[a_i, b_i]$ and so $a_i \leq a_{i+1} \leq b_{i+1} \leq b_i$ holds. Similarly, $c_i \leq c_{i+1} \leq d_{i+1} \leq d_i$. Since, by the induction hypothesis, $0 < a_i$ and $b_i \leq c_i$, we have P1.

To get P2, first note that, by the definition of depth for terms involving rec, D is greater than any number k such that $\text{rec}^k_{\sigma\tau} \in \mathcal{T}^\Phi_D$. If f is any function of the language other than $\min_\varphi$ and $f(\vec{\tau}) \in \mathcal{T}^\Phi_i$ then

$$|\text{val}(\vec{\tau})| \in \{0\} \cup [\frac{1}{c_i}, c_i] \text{ implies } |\text{val}(f(\vec{\tau}))| \in \{0\} \cup [\frac{1}{2^{c_i}_D}, 2^{c_i}_D] \cup \{\uparrow\}.$$

By the definition of the f_n's, and noting that $c_{i+1} = f^j_{D-(i+1)}(c_i)$ for some $i < D$ and $j \geq 1$, it follows that

$$c_{i+1} \geq 2^{c_i}_D.$$

So by P2, in the induction hypothesis, if $f(\vec{\tau}) \in \mathcal{T}^\Phi_{i+1}$, then $|\text{val}(\vec{\tau})| \in \{0\} \cup [\frac{1}{c_i}, c_i]$, so, if f is not $\min_\varphi$, then

$$|\text{val}(f(\vec{\tau}))| \in \{0\} \cup [\frac{1}{c_{i+1}}, c_{i+1}],$$

and hence $\frac{1}{|\text{val}(f(\vec{\tau}))|}, |\text{val}(f(\vec{\tau}))| \leq c_{i+1}$. If f is $\min_\varphi$ then, by definition, $|\text{val}(f(\vec{\tau}))| \leq d_{i+1}$; since $(c_{i+1}, d_{i+1}]$ was selected to have empty intersection with S, we have $|\text{val}(f(\vec{\tau}))| \leq c_{i+1}$.

For P2, it remains to show that $|\text{val}(f(\vec{\tau}))| \notin (a_{i+1}, b_{i+1})$, but that is a direct consequence of our choice of a_{i+1} and b_{i+1}; i.e., they were chosen so that the interval $(a_{i+1}, b_{i+1}]$ has empty intersection with S. This completes the proof that P2 holds with i in place of $i+1$.

Conditions P3 and P4 have similar proofs. They follow from the fact that a_{i+1} and b_{i+1} were chosen to be sufficiently far away from a_i and b_i, respectively. Since $j \geq 1$ in (**), we know that $a_{i+1} \geq f_{D-(i+1)}(a_i)$; by the definition of the f_n's, we clearly have P3. And since $j \leq 2t_D + 1$, we have $f_{D-(i+1)}(b_{i+1}) \leq f^T_{D-(i+1)}(a_i) = b_i$, and P4 readily follows.

P5 and P6 are immediate from (**) and the corresponding equations that are used to define c_{i+1} and d_{i+1}.

Step 4: Finally we will show that the above construction is finitistic, and explain how this gives a PRA proof of the consistency of ERNA. See [10] for a discussion of the generally recognized fact that PRA captures Hilbert's notion of finitary. The construction, including the val-assignment, only involved computing the functions of ERNA on rational numbers. The only such function that is not primitive recursively computable is $\min_\varphi$, but the version of this that need to be computed in the construction incorporated search bounds that made it computable. Note that in order to compute these versions of $\min_\varphi$ we needed to evaluate φ with rational number instantiations of its variables. Again this can be done primitive recursively since φ does not contain occurrences of min. Also, in the case of the external minimum, we relied on the fact that we had a way of computing truth values for atomic formulas involving Inf.

The sequences a_i, b_i, c_i, and d_i were defined using the hierarchy of functions f_n. The function f_0 was defined as a fixed finite iteration of exponentiation base 2; this is defined in PRA, and for $n > 0$, the f_n's are obtained by iterating the previous function finitely many times. Clearly, each f_n is bounded by a fixed finite iteration of exponentiation base 2, so not only are these functions definable in PRA, they are definable in elementary recursive arithmetic. That is, for each finite n, f_n is definable in elementary recursive arithmetic, yet elementary recursive arithmetic does not prove a formalization of the statement "for all n, f_n is definable". So the consistency proof cannot be carried out in elementary recursive arithmetic. For a detailed account of elementary recursive arithmetic see [6].

Also, it is straightforward to see that the additional details of dividing the intervals into subintervals and applying the pigeon-hole principle can all be carried out in PRA. For example, the set S, defined just above $(**)$, can be coded in PRA.

In addition to noting that the construction can be carried out in PRA, we note that PRA proves a version of Herbrand's theorem. In particular, it is provable in PRA that if for any finite collection of instances of the axioms of ERNA there is a truth assignment that makes each axiom true, then ERNA is consistent. See [3] to see how theorems of this form are proved in PRA and weaker systems.

5. Finite Models

In the previous section we referred to the val assignment as a construction of a finite model. In this section we will elaborate on the meaning of that association.

Clearly, the assignment of the val function, on a finite set of terms $\mathcal{T}_D^\Phi$ yields a finite set of rational numbers, and an assignment of truth values for relations on those rational numbers. Further, since our set of terms is closed under subterms, we can view val as giving an interpretation of the function symbols of the language. Of course, typically, val will not interpret a function to be defined on all of the elements of the universe of the finite model; i.e., the functions will not be total in the model. None the less, we can view these as finite models of the relational structure of the language.

The existence of such finite models, corresponding to sets of terms $\mathcal{T}_D^\Phi$, for arbitrary positive integers D, implies the consistency of the theory (that is Herbrand's theorem). The fact that this can be shown finitisticly is the essence of our finitary consistency proof. However, it is useful to view this from an infinitary point of view. In particular, consider the set of all such finite models ordered by substructure; this ordering yields a tree. The consistency proof of the previous section shows that this tree is infinite. Also, since there are only finitely many terms of a given depth, we can structure the tree so that it is finitely branching. König's lemma then implies the tree has an infinite branch. It is straightforward to see that the union of all of the finite models along an infinite branch is an infinite model of the theory (terms of arbitrary depth are interpreted).

Also, it is clear that the finite models along any infinite branch are substructures (as relational models) of the infinite model obtained as the union. These facts inspire the following questions:

1. To what extent can we find finite models that are substructures of infinite models?
2. Does the algorithm, of the previous section, for assigning val, yield such models?

The answers to these questions is primarily negative, and we will explain how so in this section. In the following sections we consider related questions, with the focus of the final section on positive partial results.

Although the construction of the previous section proves that there are finite models that are substructures of infinite models, there is no way to be sure that the models that are constructed are themselves substructures of infinite models. The models are produced to satisfy conditions on a fixed finite set of terms, but it may very well be the case that a given model cannot be "expanded" to interpret additional terms. We will illustrate this with an example, but first note that in the tree picture described above, it may be the case that the constructed model is an end node (leaf) of the tree. Keeping in mind that Königs lemma is non-constructive, there is no seeming contradiction in the fact that constructing substructures of infinite models is impossible.

For a simple example of a finite model that cannot be expanded to an infinite one, consider the following set of terms:

$$\mathcal{T} = \{0, 1, 1+1, \nu_0, (1+1)^{\nu_0}, (\nu_0)^{1+1}\}.$$

By setting $\text{val}(\nu_0) = 4$, and taking $\text{val}(\text{Inf}(\tau)) = \textit{true}$ if and only if $|\text{val}(\tau)| \leq 1/4$, an assignment is determined in which all axioms involving the terms in $\mathcal{T}$ are assigned *true*. However, in this assignment,

$$2^{\nu_0} = (\nu_0)^2,$$

which can be refuted by the axioms, and hence this finite model is not a substructure of any infinite model. Note that in order to refute this statement, additional terms must be brought in, and once enough additional terms are brought in the above val-assignment cannot be used.

6. The Isomorphism Theorem

In this section we state and prove the isomorphism theorem mentioned in the introduction. To amplify what was said earlier, the theorem supports the strong physical intuition that no experiment can successfully test when a physical quantity, such as time or density, is, when represented by a mathematical function, necessarily mathematically continuous or discrete. In particular, the theorem shows that the finite set of terms of a physical problem and its solution, expressible in the language of ERNA, must have, isomorphic to the standard interpretation of these terms, a finite set of rational numbers.

The theorem also supports another intuition. The powerful results of classical analysis, so widely used in physics and other sciences, do not reflect directly actual properties of physical quantities, but rather efficient computational schemes for analyzing and predicting natural phenomena. In summary, the classical representation of many physical quantities as functions having strong smoothness properties is not something given in nature, but is computationally convenient, in the sense that one frame of reference is selected rather than another strictly for computational purposes.

The following definition is used in the statement of the isomorphism theorem. A model of ERNA is **reasonably sound** if whenever $\exists n\phi$ is satisfied in the natural numbers of the model (the interpretation of $\mathcal{N}$ in the model), where ϕ is quantifier free and does not involve Inf, ϵ_0, or ν_0, then $\exists n\varphi$ is also true in the standard natural numbers. This is a version of what is sometimes called 1-consistency. By "standard natural numbers" we mean the natural numbers of the "meta-theory"; that is, in the theory in which all arguments of this paper take place. The standard natural numbers are isomorphic to the subset of the interpretation of $\mathcal{N}$ that consists of numbers that can be represented by (finite) terms of the form $0+1+1+\cdots+1$. Note that there may be elements n of a model $\mathcal{M}$ of ERNA, such that, *in the model*, n is finite and n is a natural number, i.e.,

$$(*)\quad \mathcal{M} \models n \not\approx \infty \wedge \mathcal{N}(n),$$

but n is not *actually* finite (in the sense that it cannot be represented by a finite term of the form $0+1+1+\cdots+1$). Elements n of $\mathcal{M}$ that satisfy $(*)$ are called **$\mathcal{M}$-finite**. Sometimes we will use the terminology **truly finite** or **truly standard** for the finite elements of *the standard model*. We assume a similar definition for **standard rational number**.

Theorem 6.1. Isomorphism Theorem *Let $\mathcal{M}$ be a reasonably sound model of nonstandard analysis, and let $\mathcal{T}$ be a finite set of terms in the language of* ERNA, *closed under subterms. Then there are arbitrarily large (truly) finite natural numbers b, such that there is an isomorphism f from $\mathcal{T}^{\mathcal{M}} =_{\text{def}} \{\tau^{\mathcal{M}} : \tau \in \mathcal{T}\}$, the $\mathcal{M}$ interpretations of the elements of $\mathcal{T}$, to a finite subset of the (truly) standard rationals that satisfies the following:*

1. $f(g(\tau_1, \ldots, \tau_l)) = g(f(\tau_1), \ldots, f(\tau_l))$, *where* g *is any function symbol of the language other than* ν_0 *and* ϵ_0.
2. $f(\nu_0) = n_0$ *and* $f(\epsilon_0) = 1/n_0$, *for some* $n_0 \geq b$.
3. $\mathrm{Inf}(\tau)$ *holds in* $\mathcal{M}$ *if and only if* $|f(\tau)| \leq \frac{1}{b}$.
4. $\mathcal{N}(\tau)$ *holds in* $\mathcal{M}$ *if and only if* $f(\tau)$ *is a natural number.*
5. $\tau \leq \sigma$ *holds in* $\mathcal{M}$ *if and only if* $f(\tau) \leq f(\sigma)$.

Corollary 6.2. *For* $\mathcal{M}$ *and* $\mathcal{T}$ *as in the statement of the theorem, and let* f, f' *be two isomorphisms given by the theorem. Then the finite models determined by* f *and* f' *are isomorphic.*

The corollary follows readily from the transitivity of the isomorphism relation.

Proof of Theorem:

To prove the theorem we will first show that the rationals are closed under all of the functions of ERNA; this is spelled out precisely in the second lemma below. From that lemma we will conclude that all closed terms, that are defined in a model, are interpreted by rationals, in that model. This fact will allow us to show that the assertion of properties (1–5), for the terms in $\mathcal{T}$, can be expressed by a formula of the form $\exists n\varphi$, where φ is quantifier-free. The reasonable soundness condition thus implies that this formula is realized in the standard rationals. This implies that all of the terms, together with b and n_0, can be interpreted by standard rationals in such a way that (1–5) are satisfied. This implies the theorem. This is worked out in greater detail below.

Lemma 6.3. *In* ERNA *the set of natural numbers is closed under* $|\ |$, $+$, *and* $\cdot$.

This lemma is used mainly to prove the next lemma, Lemma 6.4, which plays an important role in the proof of the theorem, as was explained above.

Proof of Lemma 6.3:

That $\mathcal{N}$ is closed under $|\ |$ follows directly from the field axiom that asserts $x \geq 0 \rightarrow |x| = x$ and axiom (N3). To see that $\mathcal{N}$ is closed under $+$ and $\cdot$, it is enough to carry out the usual proof by induction, which can be carried out in ERNA. In particular, to get

$$\mathcal{N}(x) \wedge \mathcal{N}(y) \rightarrow \mathcal{N}(x+y),$$

fix x such that $\mathcal{N}(x)$, and suppose, for a contradiction, that for some y, $\mathcal{N}(y)$ and $\neg\mathcal{N}(x+y)$. Use (IM) to get the least y such that $\mathcal{N}(y)$ and $\neg\mathcal{N}(x+y)$. Note that $y \neq 0$ since, by a field axiom, $x + 0 = x$ and we are assuming $\mathcal{N}(x)$. Using (N2), we have $\mathcal{N}(y-1)$. By the leastness of y, and using some field axioms, $\mathcal{N}(x + (y-1))$. Using (N1), and some field axioms, we have $\mathcal{N}(x+y)$, the desired contradiction. The proof for multiplication is similar. This concludes the proof of Lemma 6.3.

It will be useful to define "x is rational" by the formula:

$$\exists m \exists n (n \neq 0 \wedge |x| = \frac{m}{n}).$$

Recall the convention, Clause 3 of Section 2.2, of using m and n to range over elements of $\mathcal{N}$.

Lemma 6.4. *In* ERNA *the rationals are closed under all "defined" functions in the language of* ERNA. *In particular, for any term, $\tau(\vec{x})$, in the language of* ERNA, *where $\vec{x}$ includes all free variables of the term,* ERNA *proves that if $\tau(\vec{x}) \neq \uparrow$ and "$x_1, \ldots, x_l$ are rational" then "$\tau(\vec{x})$ is rational".*

Proof of Lemma 6.4:

The proof goes by induction on the construction of τ (more precisely, by induction on the depth of τ, where depth is defined in Section 4). Clearly, using (N1), (N4), and (I8), ERNA proves that 0, 1, ν_0, and ϵ_0 are rational.

Using the field axioms, it is straightforward to show, using Lemma 6.3, that the absolute value, sum, difference, and product of rationals is rational. Also, it is nearly immediate to see that the application of $\lceil\ \rceil$, $\pi_{l,i}$, or $\min_\varphi$ to rationals is rational (note the default value 0 for $\min_\varphi$).

We have left to show that the rationals are closed under exponentiation (to integer powers), and under functions defined by recursion using rec. The two cases are similar; they both use induction. First, suppose, for a contradiction, that for some x and y, $x^y \neq \uparrow$ and $\neg\mathcal{N}(x^y)$. Using (IM), suppose $\mathcal{N}(x)$ and suppose y is least such that $\mathcal{N}(y)$ and $\neg\mathcal{N}(x^y)$. Since $x^0 = 1$, we know that $y \neq 0$, so, by (N2), $\mathcal{N}(y-1)$, and, by the leastness of y, $\mathcal{N}(x^{y-1})$. By (E), $x^y = x^{y-1} \cdot x$, but then, by Lemma 6.3, $\mathcal{N}(x^y)$, a contradiction. The proof for the more general case of terms involving rec, is very similar. This concludes the proof of Lemma 6.4.

Now we will apply Lemma 6.4 to prove the isomorphism theorem. Let $\mathcal{T}$ be a finite set of terms in the language of ERNA, closed under subterms. Suppose

$$\mathcal{T} = \{\tau_1, \ldots, \tau_k\}.$$

By Lemma 6.4, each term in $\mathcal{T}$ is realized by a rational number. We will use a standard pairing function to express this fact as well as the statement that all of the relationships between the terms in $\mathcal{T}$ as expressed in (1–5) in the statement of the theorem, hold. In particular, we coded the pair (m, n), using the standard pairing function

$$(m, n) \mapsto \frac{(n+m)(n+m+1)}{2} + m.$$

Using this coding of pairs, we can code sequences of fixed finite length (this will be sufficient for our purposes; of course, ERNA has coding machinery to handle the coding of sequences, of natural numbers, of arbitrary length). Also, this pairing function has decoding functions that are definable in ERNA. So, there is a formula of ERNA that expresses "there exists an n such that

- n codes a sequence $(b, n_0, m_1, n_1, \dots, m_k, n_k)$
- For $i = 1, \dots, k$,

$$\frac{m_i}{n_i} = \delta(\tau_i) f\left(\frac{m_{i_1}}{n_{i_1}}, \dots, \frac{m_{i_l}}{n_{i_l}}\right),$$

  whenever $\tau_i = f(\tau_{i_1}, \dots, \tau_{i_l})$, f is not ν_0 or ϵ_0, and $\delta(\tau_i)$ is 1 or -1 according to whether $\tau_i \geq 0$ or $\tau_i < 0$, respectively, in $\mathcal{M}$.
- $\frac{m_i}{n_i} \leq \frac{m_j}{n_j}$ if and only if $\tau_i \leq \tau_j$ is true in $\mathcal{M}$.
- $\mathcal{N}(\frac{m_i}{n_i})$ if and only if $\mathcal{N}(\tau_i)$ is true in $\mathcal{M}$.
- $\frac{m_i}{n_i} < \frac{1}{b}$ if and only if $\mathrm{Inf}(\tau_i)$ is true in $\mathcal{M}$.

The formula just described is a formula that is of the form $\exists n\varphi$ where φ is quantifier free. Clearly, it is true in $\mathcal{M}$, and so, by the reasonable soundness of $\mathcal{M}$, it is true in the standard rationals. That means that n is realized by a truly standard natural number. Note further that we can incorporate into the above formula a condition $b > \overline{B}$, where $\overline{B}$ is any closed term with the truly finite value B. The formula is still true in $\mathcal{M}$, and hence, realized in the standard rationals. This implies that there are arbitrarily large finite values of b that work. This concludes the proof of the Isomorphism Theorem.

7. Constructing Finite Models

In Section 4 we introduced a construction for obtaining finite models of ERNA, but in Section 5 we showed that the construction does not necessarily lead to a model that is a substructure of an infinite model. In Section 6, the isomorphism theorem asserts the existence of models that are substructures of reasonably sound models. In light of the discussion in Section 5, we cannot expect to construct such finite substructures, and in a moment we will give a short proof of that claim. However, we note that in some cases one can find such models, and we will illustrate this with an example.

A construction of finite models that were substructures of reasonably sound models would entail a decision procedure for formulas of the form $\exists n\varphi$ where φ is a quantifier free formula of ERNA that does not involve min, Inf, ϵ_0, or ν_0. To see that this is the case, note that by constructing an interpretation, n_φ, of a term of the form $\min_\varphi(\vec{m})$, we could then test the truth of $\exists n\varphi$ by checking to see if $\varphi(n_\varphi, \vec{m})$ holds. By the reasonable soundness condition, we would be evaluating the truth of $\exists n\varphi$ in the standard model. But we cannot expect to evaluate such formulas in general since that would entail a decision procedure for arbitrary recursively enumerable sets. This follows from the fact that any recursively enumerable set can be represented in the form

$$\{m : \exists n\varphi(m, n)\}$$

for the appropriate quantifier free formula φ in the language of ERNA with no occurrences of $\min_\varphi$, ϵ_0, or ν_0. For some of the details of the argument needed to show this see [6].

As an example of a construction of a finite model, we consider the terms that are involved in the analysis of the physical problem of a freely falling body, which involves a very simple differential equation. We take as very restricted but computationally simple physical premises the antecedent of the following conditional statement, which we can then prove in ERNA.

$$(\forall t)(1 \leq t \leq 2 \ \wedge \ \ddot{x} \approx 32 \ \wedge \ x(1) = 0 \ \wedge \ \dot{x}(1) = 0 \rightarrow x(t) \approx 16t^2 - 32t + 16)).$$

Replacing the derivatives as defined in [9] we get:

$$(\forall t)(1 \leq t \leq 2 \ \wedge \ \frac{\frac{x(t+2\epsilon_0)-x(t+\epsilon_0)}{\epsilon_0} - \frac{x(t+\epsilon_0)-x(t)}{\epsilon_0}}{\epsilon_0} \approx 32 \wedge$$

$$x(1) = 0 \wedge \frac{x(1+\epsilon_0) - x(1)}{\epsilon_0} \approx 0 \rightarrow x(t) \approx 16t^2 - 32t + 16).$$

Negating the immediately preceding result, we obtain an existential statement, and we eliminate the existential quantifier with a constant c, to obtain

$$(*) \quad 1 \leq c \leq 2 \ \wedge \ \frac{x(c+2\epsilon_0) - 2x(c+\epsilon_0) + x(c)}{\epsilon_0^2} \approx 32$$

$$\wedge x(0) = 0 \ \wedge \ \frac{x(1+\epsilon_0) - x(1)}{\epsilon_0} \approx 0 \ \wedge \ x(c) \not\approx 16c^2 - 32c + 16$$

Looking at $(*)$ we have the following:

- Finite non-infinitesimal terms: $16c^2$, 32, $32c$, $32(c+\epsilon_0)$, $32(c+2\epsilon_0)c$, c^2, $c+\epsilon_0$, $c+2\epsilon_0$, 16, $16(c+\epsilon_0)$, $16(c+\epsilon_0)^2$, $16(c+2\epsilon_0)$, $16(c+2\epsilon_0)^2$,

where $x(c)$, $x(c+\epsilon_0)$, $x(c+2\epsilon_0)$, $2x(c+\epsilon_0)$, $x(1+\epsilon_0)$, $x(1)$ are replaced using $x(t) \approx 16t^2 - 32c + 16$, which we can prove is monotonically increasing for $t \geq 1$.

- Infinitesimal terms in $(*)$: ϵ_0, ϵ_0^2, $16\epsilon_0$, $32\epsilon_0$, $64\epsilon_0$, $32c\epsilon_0$, $64c\epsilon_0$, $16\epsilon_0^2$, $64\epsilon_0^2$.

Since $c \geq 1$, $64c\epsilon_0$ is obviously the maximal infinitesimal term.
Also, $16(c+2\epsilon_0)^2$ is the maximal finite term. Noting that $c \leq 2$, this maximal term is bounded from above by 65 for ϵ_0 small, and so we take $b = 2^6 + 1 = 65$.

Then $64c\epsilon_0 \leq 128\epsilon_0 \leq \frac{1}{b}$, so

$$\epsilon_0 \leq \frac{1}{128b} = \frac{1}{128 \cdot 65} = \frac{1}{2^7(2^6+1)}.$$

so the n_0 required for $c = 2$ must satisfy:

$$128 \cdot 65 \leq n_0,$$

and we assume equality.

To show what happens as the upper bound of t increases, we give approximate values for b and n_0. For $t = 1$, the case just computed, take $b = 2^6$ and $n_0 = 2^{13}$ as good approximations. Similarly, for $t = 2^8$, $b = 2^{20}$ and $n_0 = 2^{34}$, and for $t = 2^{20}$, $b = 2^{44}$ and $n_0 = 2^{70}$, so in this example, convergence of the natural sequence of finite models to the infinite model is rapid.

References

[1] Rolando Chuaqui and Patrick Suppes. An equational deductive system for the integral calculus. In P. Martin-Löf and G. Mints (Eds.) *Lecture Notes in Computer Science, Proceedings of COLOG-88, International Conference on Computer Logic, Tallinn, USSR.* Berlin Heidelberg, Springer Verlag. pp. 25–49.

[2] Rolando Chuaqui and Patrick Suppes. Free-variable axiomatic foundations of infinitesimal analysis: a fragment with finitary consistency proof. *JSL*, 60(1):122–159, 1995.

[3] P. Hájek and Pavel Pudlák. *Metamathematics of First Order Arithmetic.* Springer-Verlag, 1993.

[4] A. E. Hurd and P. A. Loeb. *An Introduction to Nonstandard Real Analysis.* Academic Press, 1985.

[5] Jeff B. Paris and L.A.S. Kirby. Initial segments of models of Peano's axioms. In *Springer-Verlag Lecture Notes in Mathematics*, volume 619, pages 211–226. Springer-Verlag, 1977.

[6] H. E. Rose. *Subrecursion: Functions and Hierarchies.* Clarendon Press, 1984.

[7] Richard Sommer *Transfinite Induction and Hierarchies Generated by Transfinite Recursion within Peano Arithmetic.* PhD thesis, U. C. Berkeley, 1990.

[8] Richard Sommer Transfinite Induction in Peano Arithmetic. To appear in *Annals of Pure and Applied Logic*, 1996.

[9] Patrick Suppes and Rolando Chuaqui. A finitarily consistent free-variable positive fragment of infinitesimal analysis. *Proceedings of the IX Latin American Symposium on Mathematical Logic,* Notas de Mathematica, Universidad Nacional del Sur, Bahia Blanca, Argentina, 1995.

[10] G. Takeuti. *Proof Theory.* North Holland, 1987.

STANFORD UNIVERSITY