Reprinted from JOURNAL OF MATHEMATICAL PSYCHOLOGY

Vol 41, No 1, March 1997
Printed in Belgium

Dispensing with the Continuum

Richard Sommer and Patrick Suppes

CSLI, Stanford University, Stanford, California 94305-4115

We present a constructive system of nonstandard analysis, called elementary recursive nonstandard analysis (ERNA), as a viable foundation for the mathematics that is used in the empirical sciences; the viability is demonstrated by showing that a version of the standard existence theorem for first-order ordinary differential equations is provable in ERNA. We demonstrate the constructive character of ERNA by showing that it has a finitary consistency proof Through the consistency proof one can make the observation that ERNA has models in which every element is a (possibly nonstandard) rational number; hence properties special to the continuum are not used in ERNA. Also, we will show how the consistency proof leads to the construction of finite models in the standard rationals. Then we give an isomorphism theorem stating that the interpretation of any finite set of ERNA-terms can be mapped, isomorphically, onto a finite set of standard rationals. Additionally, we discuss to what extent such isomorphisms are constructive and how the isomorphism theorem lends further support to the thesis that the continuum is dispensable in the mathematics that is used in the empirical sciences. © 1997 Academic Press

1. INTRODUCTION

The reintroduction of infinitesimals by Abraham Robinson in the 1960s revived ideas that were dormant for well over 100 years but that had earlier played an important role in geometry in the 1500s and 1600s and later in the development of the calculus. It is still the case that many mathematicians are skeptical about the desirability of explicitly introducing infinitesimals and prefer to stay with the classical foundations of analysis developed by a number of mathematicians, but especially Dedekind and Weierstrass in the 19th century.

Ironically, this arithmetization of analysis was itself severely criticized by leading mathematicians of the time, such as Hankel and du Bois-Reymond, and later, Veronese. The central objection was the artificiality of the construction of irrational numbers without the traditionally important concept of magnitude. The second related objection was against the set-theoretical construction of the continuum. If a continuous line is composed of elements, the elements must be extended, not dimensionless points. Veronese was particularly articulate on this issue. And those elements were, of course, infinitesimals of one kind or another. Moreover, infinitesimals, like magnitudes in general, were geometrical in origin. Their use by Simon Stevin, Kepler, Cavalieri, and others predates Newton and Leibniz by many decades.

Reprint requests and correspondence should be addressed to. Richard Sommer, CSLI, Stanford University, Stanford, CA 94305-4115

In this paper, we want to give a sense of some of the things that can be done with infinitesimals, especially within a constructive system which we call elementary recursive nonstandard analysis (ERNA). We claim this system is sufficient to provide a foundation for a significant part of the mathematics used in the sciences.

We begin with a brief informal description of nonstandard number systems in general. Results in mathematical logic (in particular, the compactness theorem) imply that it is consistent with the theory of the natural numbers to have infinite numbers, i.e., natural numbers that are greater than all integers represented by terms of the form $0+1+1+1+\cdots+1$. Such numbers are called nonstandard natural numbers, e.g.,

$$\underbrace{0 \quad 1 \quad 2 \quad 3 \quad 4 \quad \cdots}_{\text{standard numbers}} \qquad \underbrace{\cdots \quad \nu-1 \quad \nu \quad \nu+1 \quad \cdots}_{\text{nonstandard numbers}}$$

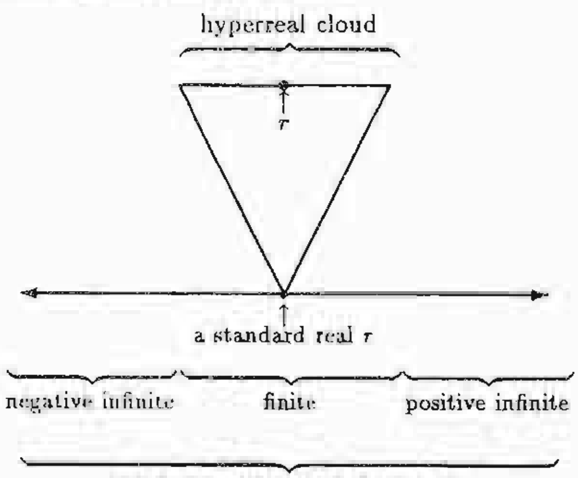
For any given set of true statements of number theory, there are extensions of the natural numbers which include nonstandard numbers and satisfy all statements in the given set. In fact, in each case, there are infinitely many choices for the nonstandard part of the extension. Let N^* denote an extension of the natural numbers that includes nonstandard numbers. From N^* we can define an extension Z^* of the integers Z that includes negative nonstandard integers, and Z^* can be used to obtain an extension Q^* of the rationals Q which is to include all fractions of (possibly nonstandard) integers with nonzero denominators. Then Q^* can be used to obtain an extension R^* of the real numbers; yet, the way in which we go from Q^* to R^* is not at all straightforward (as it is for going from, say, Z^* to Q^*). However, versions of the usual constructions of the reals will work in this setting (for example, constructions based on Dedekind cuts, Cauchy sequences, or decimal expansions ad in each case, care has to be taken to have these constructions be formulated in terms of *internal* sets (a concept to be defined later), rather than in terms of all sets).

0022-2496/97 $25 00

For the remainder of this section we fix particular extensions Q^* and R^* as described above, and we will call these sets the *hyperrationals* and the *hyperreals*, respectively. The reader should be cautious of the fact that there is no unique choice for Q^* and R^*, so one needs to keep in mind that the use of the terminology *the* hyperrationals and *the* hyperreals is a result of fixing Q^* and R^* from among infinitely many possibilities. This section is meant to be primarily motivational; later sections will focus mainly on the syntactic properties of ERNA and will not rely on any properties of Q^* and R^* presented here. In particular, specific properties, such as exactly how much of the true theory of Q and R are satisfied in Q^* and R^*, are not important for this motivational discussion.

An element of Q^* or R^* is *infinite* if, in absolute value, it is greater than a nonstandard natural number, and a number is *infinitesimal* if its reciprocal is infinite, or if the number is zero. If $x - y$ is an infinitesimal we say *x is approximately equal to y* or *x is infinitely close to y*, and we write $x \approx y$. The relation $\approx$ is an equivalence relation.

We refer to an $\approx$-equivalence class as a *hyperreal cloud*. Note that each standard real number sits in a unique hyperreal cloud. The following picture is useful for developing an intuition of the hyperreal numbers:



THE HYPERREAL LINE R^*

Note that each hyperreal number can be identified with a decimal expansion that includes decimal places corresponding to both the standard and nonstandard natural numbers; in particular all hyperreals are of the form

$$z.d_1\,d_2\,d_3\ldots d_{\nu-1}\,d_\nu\,d_{\nu+1}\ldots, \qquad (*)$$

where z is a (possibly nonstandard) integer and the d_i's are digits amongst 0, 1, 2, ..., 9. Any pair of hyperreals that have the same value of z and the same decimal expansion through all standard decimal places (i.e., all d_i's are the same if i is standard) differ by at most an infinitesimal; i.e., they are approximately equal. It is clear that every $\approx$-equivalence class contains rational numbers; one way to see this is to consider the truncation of the decimal expansion at a fixed nonstandard decimal place:

$$z.d_1\,d_2\,d_3\ldots d_{\nu-1}\,d_\nu. \qquad (**)$$

The hyperreals given by $(*)$ and $(**)$ lie in the same hyperreal cloud. Note that while $\sqrt{2}$ is not in Q^*, there may be elements of Q^* that have the same standard decimal places as $\sqrt{2}$.

Part of the point of ERNA is its use in showing that mathematics carried out for scientific purposes does not rely on the existence of a completed continuum. The structure of the hyperrationals alone is rich enough to carry out such mathematics. Standard definitions of continuity, differentiability, and other notions from analysis rely in an essential way on the existence of a completed continuum; standard analysis takes as basic a completeness axiom such as the least upper bound principle. On the other hand, the corresponding definitions in nonstandard analysis do not rely on such principles. By trading in the completeness axioms for axioms asserting the existence of infinitesimals, we end up with a system that is actually more constructive and, in many ways, better matches certain geometric intuitions about the number line.

A detailed description of the axioms of our system is given in Section 2. Fundamentally, what is done with ERNA is to develop Q^* as a weak system of analysis. We claim that within ERNA we can do much of the mathematics that is used in the empirical sciences, in particular, those parts that are considered constructive, and we are in fact skeptical that anything other than constructive results are really required for scientific work concerned with any kind of experimental or empirical results. As a way to demonstrate the potential of ERNA to carry out mathematics that is applied in the sciences, in Section 3 we discuss how ERNA is sufficient to prove an existence theorem for first-order differential equations. We note that there is nothing special about this existence theorem that makes it especially adaptable to ERNA; many other classical theorems that are used in mathematical practice also have versions provable in ERNA.

In Section 4 we describe a finitary consistency proof for ERNA. On the one hand, this demonstrates the constructive nature of ERNA, but, more importantly for the main thesis of this paper, the consistency proof allows us to see that ERNA has models in which every element is a (possibly nonstandard) rational number. In other words, the statement "every element is a rational number," as expressed in the language of ERNA, is consistent with ERNA. From this we know that proofs in ERNA do not rely on the completeness property of the continuum, so, using the results of Section 3, we conclude that a significant portion of the mathematics that is used in the sciences does not rely on special properties of the continuum.

The consistency proof of Section 4 can be understood as the description of a method for constructing finite models, in subsets of the standard rationals, of finite sets of ERNA terms; this is discussed further in Section 5 and is used to lead in to our isomorphism theorem which asserts the existence of such models for any finite collection of ERNA terms. The models referred to in the isomorphism theorem are isomorphic to finite substructures of nonstandard extensions of the continuum. In Section 6 we tie together themes of Sections 3, 4, and 5. In particular, we argue that much of the mathematics of the empirical sciences can be modeled in finite sets of rational numbers; again, we have support for the dispensability of the continuum in mathematical practice.

To illustrate the nature of some of our results, in particular the isomorphism theorem, presently we will give a simple example that is a demonstration of how the continuum is not empirically needed mathematically. The isomorphism theorem presented in Section 5 is similar, but it is a much stronger result. We will make a claim using only what is called the hyperfinite grid G^* which consists of equally spaced points (hyperreal numbers) such that the distance between each of two adjacent points is a fixed infinitesimal (so $G^* \subset R^*$). Thus each point has a unique successor and a unique predecessor; this is in contradiction to the standard properties even of the rational line, let alone the ordinary continuum.

Let R denote the standard real numbers. Note that neither R nor G^* is a subset of the other, but both are subsets of the set R^* of hyperreal numbers. For any finite x in R^*, we get to R, so to speak, by taking the standard part of x, denoted $\mathrm{std}(x)$; the real number $\mathrm{std}(x)$ is the unique standard number that is infinitely close to x ($\mathrm{std}(x)$ is only defined when x is finite). As a definition we can specify that, if x is finite, $\mathrm{std}(x) - x$ is infinitesimal and $\mathrm{std}(x)$ is in R. Now suppose ψ is a function from R to R, and suppose φ is a function from G^* to G^*. Then we say that ϕ and ψ are *empirically indistinguishable* if and only if for any real number x in R and any number q in G^* such that the difference between x and q is infinitesimal, the difference between $\psi(x)$ and $\phi(q)$ is infinitesimal. The intuitive picture here is quite clear. The hyperfinite grid provides an equally spaced partition of the hyperreal line, so any real number is within an infinitesimal of a point on the hyperfinite grid. In fact, it does not take very much mathematical apparatus to prove the following.

PROPOSITION. *Let ψ be a function from R to R. Then there exists a function ϕ from G^* to G^* empirically indistinguishable from ψ.*

To obtain ϕ from ψ, for q finite, set $\phi(q) = \psi(\mathrm{std}(q))$, and set $\phi(q)$ arbitrarily if q is infinite. If x, in R, is infinitesimally close to q then $x = \mathrm{std}(q)$ and so $\phi(q) = \psi(x)$; so clearly they are infinitesimally close.

To our mind, the concept of empirical indistinguishability we have introduced for functions expresses a very fundamental insight about empirical science. We cannot conceive of any experiments that would differentiate between two concepts whose quantitative expression differed only by infinitesimal amounts. On the other hand, in the realm of pure mathematics it is easy enough to understand and have a feeling for the difference between functions defined only on the hyperfinite grid and those defined on the standard continuum, or even the richer hyperreal line.

For those who would like to see an example of a purely mathematical application of the hyperfinite grid that also demonstrates the kind of usefulness that is important in science, we consider the lifting method to prove results ordinarily formulated for the continuum. The approach goes like this. First, the mathematical objects that are given are "lifted" to approximations on the hyperfinite grid. Second, elementary computations are then made on the hyperfinite grid to construct, for example, a solution or some other object. This new object or solution is then taken back to the standard real line in order to obtain a solution to the original problem. In particular, one can prove existence theorems for stochastic differential equations in this fashion by solving the corresponding stochastic difference equations on the hyperfinite grid and then taking standard parts to obtain a solution of the original stochastic differential equation. For examples of this strategy at work, see Albeverio *et al.* (1986).

While the above example of translating from the continuum to the hyperfinite gives support to the dispensability of continuity in the empirical sciences, one may object on the basis that the hyperfinite grid is but an abstraction that contains the continuum in disguise. The aim of this paper is to improve upon this type of result in a significant way; in particular, in such a way as to avoid the just-mentioned objection.

2. THE AXIOMS OF ERNA

In this section we will give an informal description of the axioms of ERNA; for a more formal description see Sommer and Suppes (1996). ERNA is formulated in a first-order language with various function and relation symbols that are described below. The variables are intended to range over an extension of the rationals that includes infinitesimals.

Natural number axioms. ERNA contains a symbol $\mathcal{N}$ that is intended to denote the set of natural numbers, and ERNA includes the basic axioms for the set of natural numbers; in particular, the axioms expressing 0 *is the least natural number*, *the natural numbers are closed under successor*, *and the nonzero natural numbers have natural number predecessors*. The set of natural numbers will include "infinite" numbers—this is explained in more detail below.

Infinitesimal axioms. ERNA contains a symbol *Inf* that is intended to denote the set of *infinitesimals*. An infinitesimal number is a number that is extremely close to zero in a sense that can be made precise as follows. Dual to the notion of infinitesimal number is that of *infinite* number. The infinite numbers are just the reciprocals of the nonzero infinitesimals, and, conversely, the nonzero infinitesimals are just the reciprocals of the infinite numbers. Numbers that are not infinite are finite. Included in the axioms for infinitesimals is an axiom asserting that *the finite numbers are closed under addition* (the sum of two finite numbers is finite); also there is an axiom implying that there is an infinite natural number; in particular, there are constant symbols ν_0 and ε_0, and axioms asserting *ν_0 is a natural number and $\varepsilon_0 = 1/\nu_0$ is infinitesimal.* From these axioms we can discern a partition of the natural numbers into the finite and the infinite, where the finite natural numbers are closed under taking successors, so the finite natural numbers will include all numbers that can be represented by terms of the form "$0 + 1 + 1 + \cdots + 1$." Dually the infinitesimals are all smaller than every number of the form $1/n$, where n is a finite positive integer; this is what is meant when we say the "infinitesimals are extremely close to zero."

The other infinitesimal axioms further support the intuitive picture alluded to above; they include closure conditions for the operations of addition and multiplication and the order relation (*if $|x| < y$ and y is infinitesimal, then x is infinitesimal*).

Field axioms. ERNA includes the axioms for an ordered field.

Exponential function axioms. ERNA includes a symbol for exponentiation to natural number powers and the recursive defining equations for this function. As will be explained below, all functions defined in the language are bounded in growth rate by a (standard) finite iteration of the exponential function. This feature of ERNA is very important for maintaining its constructive character and for allowing a finitary consistency proof. Functions that come up in computations in the empirical sciences are nearly always bounded in growth rate by iterations of the exponential function.

Recursion axioms. ERNA allows for a form of *limited definition by recursion*. In particular, ERNA allows for a function f to be defined in terms of functions g and h by the scheme

$$f(0) = g, \qquad f(n+1) = h(f(n), n),$$

where f, g, and h may involve additional arguments. But definition by recursion is allowed only in the event that the function so defined is bounded in growth rate by a (standard) finite iteration of the exponential function. This is slightly messy to formulate precisely; we refer the reader to Sommer and Suppes (1996) for such a formulation. Definition by recursion allows us to introduce functions such as factorial and summation into ERNA.

Minimum axioms. ERNA has two forms of induction on quantifier-free formulas. One is equivalent to ordinary mathematical induction for quantifier-free formulas that do not involve the symbol *Inf*. This axiom is formulated as *a minimum natural number principle* using *a minimum operator* and is called the *internal minimum axiom*. Essentially, the axiom asserts that every nonempty set of natural numbers that is quantifier-free definable without using *Inf* has a least element. It is easy to establish that this axiom is equivalent to a form of induction.

In general, formulas involving *Inf* are called *external*; formulas not involving *Inf* are *internal*. The reason for disallowing *Inf* in the above minimum axiom is that it is inconsistent to have a least infinite natural number, and such assertions, as well as other inconsistencies, would be provable if *Inf* was allowed to appear in a formula to which the minimum operator could be applied. However, ERNA does have a version of induction for formulas involving *Inf*; in this case it is formulated only as a *minimum finite natural number principle*. Essentially this axiom asserts that any nonempty set of natural numbers, that is quantifier-free definable and that has a finite element has a least element. This is called the *external minimum axiom*. In both the external and internal cases the symbol for the minimum operator is not allowed to appear in the induction formula; allowing such a "nesting" would boost the strength of the system beyond the level of allowing for a finitary consistency proof.

An important property of the axioms of ERNA is that they can all be expressed without quantifiers; this may not be immediately apparent since we have not provided formal statements of all of the axioms, but it is not too hard to establish this from the description of the axioms given above. The trickiest cases are the minimum axioms, but using a symbol $\min_\varphi$ for the minimum operator for the quantifier-free formula φ, this axiom can be expressed by asserting *if $\varphi(n)$ holds, and n denotes a natural number, then $\varphi(\min_\varphi)$ and $\min_\varphi \leqslant n$.* This can all be done in a quantifier-free way. For complete details the reader is referred to Sommer and Suppes (1996).

3. AN EXISTENCE THEOREM FOR DIFFERENTIAL EQUATIONS IN ERNA

The purpose of this section is to illustrate the potential ERNA has for proving mathematical theorems that are important in the sciences. To set the stage, we will say how the notions of continuity and differentiability are defined in ERNA. These will just be the usual definitions from nonstandard analysis.

Recall that $a \approx b$ means that $a - b$ is infinitesimal. A function f is said to be *continuous* at x if

$$x \approx y \text{ implies } f(x) \approx f(y).$$

A function f is *differentiable* at x if for all nonzero infinitesimals ε and ε',

$$\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \approx \frac{f(x+\varepsilon')-f(x)}{\varepsilon'}.$$

In the usual formulations of nonstandard analysis, the derivative of f at x is defined to be the standard part of the above expressions, but since ERNA does not have a standard part function, we have to do things a little differently. In particular, if f is differentiable at x, we define the *derivative* of f at x by

$$f'(x) = f(x+\varepsilon_0)/\varepsilon_0.$$

Note that this definition differs from the usual one by an infinitesimal.

Now we will discuss the ERNA proof of the following theorem.

EXISTENCE THEOREM FOR FIRST-ORDER ORDINARY DIFFERENTIAL EQUATIONS. *Suppose f is continuous and finite in the rectangle*

$$R = \{(t, y): |t - t_0| \leqslant a, |y - y_0| \leqslant b\}.$$

Then there is a solution to the initial value problem:

$$y' = f(y, t), \quad y(t_0) = y_0, \tag{1}$$

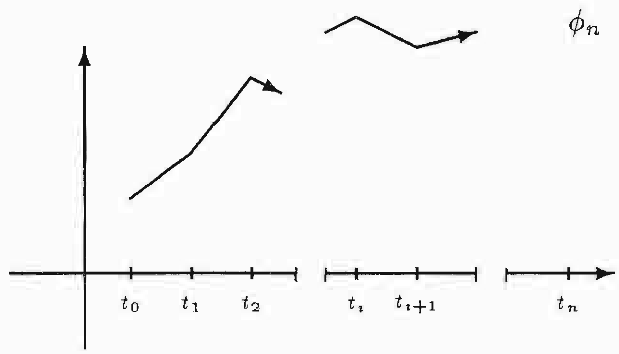
in some interval $I = \{t: |t - t_0| \leqslant h\}$.

We are being somewhat informal here; in order to state this correctly as a theorem of ERNA we would have to take into account that ERNA does not have variables ranging over functions. All functions are represented by terms. Furthermore, in ERNA we can only show the existence of an approximate solution. In other words, by "$y = \phi(t)$ is a solution to (1)," in ERNA, we mean that

$$\phi'(t) \approx f(\phi(t), t), \qquad \phi(t_0) = y_0. \tag{2}$$

For statements similar to this theorem, but correctly formulated in ERNA, see Chuaqui and Suppes (1995), Suppes and Chuaqui (1993).

The standard proofs of this theorem define a sequence of functions that are each piecewise linear on the interval I (note that h can be determined in advance); in particular, for each natural number n, a function ϕ_n is defined on I by first subdividing I into n subintervals of equal size, and then computing f at the left endpoint of each subinterval in order to determine the (constant) slope of ϕ_n in that subinterval.



This process is an application of definition by recursion; given the value of ϕ_n at the left endpoint of one subinterval, we can use f to compute the slope of ϕ_n in that subinterval, and then use the slope to get the value of ϕ_n at the left endpoint of the next subinterval. In the end, the solution to the initial value problem (1) is obtained by taking ϕ to be the limit of the ϕ_n's. Most standard textbooks in differential equations, such as Coddington and Levinson (1955), have a version of this proof.

A proof in nonstandard analysis is obtained by defining ϕ_n as above and then taking n to be an infinite natural number. The resulting function can be shown to be a solution to the initial value problem in the sense of (2). So to get a proof in nonstandard analysis, the limiting process of the standard proof is replaced by carrying out the piecewise construction on an *infinite* subdivision of I. See Hurd and Loeb (1985) for a version of this proof.

To carry out the proof in ERNA, the proof from nonstandard analysis described above is modified. In some cases the arguments are translated directly into ERNA, but in other cases careful modifications have to be made in order to have an argument that works in ERNA. However, the overall flavor of the proof retains much of the naturalness that is present in proofs in nonstandard analysis .

Even from the informal description of the axioms in the previous section it should become clear that the functions ϕ_n, defined above, are definable in ERNA; in particular, the form of the definition by recursion given in ERNA allows us to carry out the construction of ϕ_n described above. In order to see that the growth rate condition is satisfied we use the condition that f is continuous on R to conclude that f is bounded on R. This relies on the ability of ERNA to prove a version of the following lemma.

LEMMA. *A continuous function on a finite closed interval is bounded.*

We will not go into the details here, but a version of this lemma can be proved, and it will not only be used to obtain the necessary bounds on the function defined by recursion described above, but it is used in other ways in the proof of the existence theorem as well.

We note that the existence theorem above does not have special properties that make it especially adaptable to a proof in ERNA; in fact, it is presented to exhibit the versatility of ERNA in its ability to prove theorems of analysis since the proof will rely on lemmas, such as the above, that are, in themselves, of importance in analysis.

Readers familiar with the program of reverse mathematics, in particular those familiar with the results given in Simpson (1984), will recognize a discrepancy in proof-theoretic strength between the ERNA version of the existence theorem, and the version as formulated in second-order arithmetic. In the case of the ERNA version, the theorem has proof-theoretic strength at most that of elementary recursive analysis, so the second-order arithmetic version is provably stronger (see Simpson, 1984). This discrepancy can be reconciled by noting that the system in Simpson (1984) works with a base theory (RCA_0) which is already of strength greater than elementary recursive analysis.

4. CONSISTENCY PROOF

In this section we provide a brief and informal description of the finitary consistency proof given in Sommer and Suppes (1996). The consistency proof uses Herbrand's theorem which states, roughly, that the statement of consistency of a quantifier-free axiom system is equivalent to the assertion that any finite set of closed instances of the axioms can be prepositionally satisfied by a truth assignment to the atomic formulas involved. Put another way, consistency is equivalent to finding finite models for any finite set of closed instances of the axioms. An important feature of this version of consistency is that, if a constructive method for finding such finite models is found, then the resulting consistency proof is constructive.

So the way the consistency proof proceeds is by describing a constructive algorithm such that, given a finite collection of closed terms in the language of ERNA, a finite model is produced. This finite model consists of a finite set of rational numbers; each term τ in our given finite set of terms is assigned a rational number value $\mathrm{val}(\tau)$ in our finite set of rationals. In our model, $\leqslant$ and $\mathcal{N}$ are determined by the corresponding relations on the rational numbers. For example, the atomic formula $\tau \leqslant \sigma$ is assigned "true" if and only if $\mathrm{val}(\tau) \leqslant \mathrm{val}(\sigma)$ is true in the rationals. In order to assign truth values to atomic formulas involving *Inf* a natural number b is determined that corresponds to the boundary between the finite and the infinite (in our constructed model) In particular, if $\mathrm{val}(\tau) \geqslant b$ then τ is considered infinite; i.e., $Inf(1/\tau)$ is assigned "true."

For most of the functions of ERNA the val-assignment is straightforward; in particular, the assignment is made *homomorphicly*, so, for example,

$$\mathrm{val}(\sigma + \tau) = \mathrm{val}(\sigma) + \mathrm{val}(\tau).$$

In most cases this assignment is constructive, but that will not be true for the minimum operator since the minimum operator involves an unbounded search. In order to maintain the constructive character of the model construction it is necessary to place bounds on the searches involved in computing the minimum operator. As is shown in Sommer and Suppes (1996), this can be done while achieving our goal of constructing a finite model that satisfies the axioms involved.

Additionally, care has to be taken in selecting the value of the bound b, mentioned above, and the value of the infinite natural number ν_0. In particular, these values have to be carefully selected so that the "finite" is separated from the "infinite" in the manner required by the axioms for the infinitesimals. Again, this can be done constructively as described in Sommer and Suppes (1996).

An important observation that can be expressed precisely is that the above construction is finitistic (the construction can be carried out in the system of primitive recursive arithmetic (PRA)). Furthermore, Herbrand's theorem can be carried out finitistically (i.e., it can be carried out in PRA). This allows us to give a finitary consistency proof for ERNA. A more careful analysis of the above construction allows us to establish precise bounds on the *proof-theoretic strength* of ERNA; in particular, ERNA is of the same strength as the system of elementary recursive arithmetic, a system well-known to proof theorists.

Standard techniques from proof theory allow us to use the consistency proof to establish a precise characterization of the "computational strength" of ERNA. We will not go into that in any detail, but we mention that the provably recursive functions of ERNA are exactly the elementary recursive functions. We refer the reader to Takeuti (1987) to see the connections between the consistency strength and the computational content of a theory, and we refer the reader to Rose (1984) for a description of the elementary recursive functions.

Also, by analyzing the consistency proof, it is straightforward to establish that ERNA has models in which every element is a (possibly nonstandard) rational number. In particular, since the val-function takes its values in the rationals, it is consistent with ERNA to add the assertion "every element is rational," as phrased in the language of ERNA. This shows that properties special to the continuum are not used in an essential way in a proof in ERNA. Noting the results expressed in Section 3, we have that special properties of the continuum are not needed to establish the basic existence theorem for first-order ordinary differential equations, as well as other theorems that are fundamental to the mathematics that is applied in the sciences.

5. FINITE MODELS AND THE ISOMORPHISM THEOREM

Through the description of the consistency proof in terms of finite models, a number of natural questions arise. Noting that the essential requirement of the finite models of the previous section is that they be consistent, one can ask if they are substructures of a nonstandard extension of the continuum? It turns out that, in general, they are not, and, in fact, they may not be substructures of any model of ERNA. We explain this further below.

Again rephrasing, Herbrand's theorem asserts that a quantifier-free axiom system is consistent if and only if any finite set of terms in the language can be modeled in a way that is consistent with the axioms of the system. This can be viewed as a finitary version of the completeness theorem for first-order logic which asserts that an axiom system is consistent if and only if it has a model. The models of ERNA implied to exist by the completeness theorem are infinite whereas those implied to exist by Herbrand's theorem are finite; the connection can be described as follows.

The finite models of Herbrand's theorem can be ordered in a natural way using the relation of substructure to form a finitely branching tree. The consistency proof of the previous section can be used to show that this tree is infinite. By König's lemma the tree has an infinite branch. Along this infinite branch all terms in the language of ERNA are interpreted, and so we have a full model of ERNA, in the same sense as in the completeness theorem.

However, there is no constructive way to determine if a finite model from the consistency proof lies on such an infinite branch and is hence a substructure of a full model. Furthermore, the models corresponding to these infinite branches are all countable, since they have as their universe exactly the interpretations of the terms of ERNA, so they are not extensions of the standard continuum.

The above view, based on the consistency proof, is sort of a "bottom up" point of view. We obtain an interesting result if we take a "top down" point of view. In particular, we start with a nonstandard extension of the real numbers and consider the interpretation of a finite set of terms in the language of ERNA. We then show that the interpretation of those terms, considered as a substructure of the full model, is isomorphic to a finite substructure of the standard rational numbers. This is expressed more precisely in the following theorem.

ISOMORPHISM THEOREM. *Let $\mathcal{T}$ be a finite set of terms, in the language of ERNA, closed under subterms. Then there is a finite natural number b and an isomorphism f from the interpretation of the elements of $\mathcal{T}$, in a fixed extension of the continuum, to a finite subset of the 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 or ε_0.*
2. $f(\nu_0) = n_0$ *and* $f(\varepsilon_0) = 1/n_0$, *for some* $n_0 \geqslant b$.
3. $\mathrm{Inf}(\tau)$ *holds in our fixed extension of the continuum if and only if*

$$|f(\tau)| \leqslant 1/b.$$

4. $\mathcal{N}(\tau)$ *holds in our fixed extension of the continuum if and only if $f(\tau)$ is a natural number.*
5. $\tau \leqslant \sigma$ *holds in our fixed extension of the continuum if and only if*

$$f(\tau) \leqslant f(\sigma).$$

See Sommer and Suppes (1996) for a proof of the isomorphism theorem. The fact that such isomorphisms exist demonstrates that the mathematical description of a situation, as given by the interpretations of the relevant terms and the relationships between them, may just as well take place in a finite set of rational numbers, as take place in the continuum. The isomorphism theorem says that, given a finite set of ERNA terms, there is a finite model in the rationals formally indistinguishable from the interpretation of the terms in the continuum. Hence, again we see that special properties of the continuum have no role in providing a mathematical description of a situation.

6. CONCLUDING REMARKS

We now return to the existence theorem proved in Section 3 and consider how it can be understood first in terms of the construction described in the consistency proof in Section 4 and in terms of the isomorphism theorem given in Section 5.

The proof of the existence theorem, as it is described in Section 3, holds for an arbitrary function f with the desired properties; however, formalized in ERNA, f would be instantiated by a term in the language, and the proof as described in Section 3 can be thought of as a proof scheme. Consider the following situation. We are given an instance of the initial value problem (1), where f is instantiated by an ERNA term. Furthermore, we may be given some constants $c_1, \ldots, c_k$, expressed as closed ERNA terms, and in addition to wanting to know the solution ϕ to (2), we suppose we want to know some of the values $\phi(c_i)$, when c_i is in the interval I, and whether certain relations, each of the form

$$\phi(x_i) \leqslant \phi(c_j) \qquad \text{or} \qquad \phi(x_i) \leqslant c_j,$$

hold.

The isomorphism theorem says that there is a finite set of rationals that will serve as the interpretations of the c_i's and the $\phi(c_i)$'s such that all relationships of the form (3) are correct. However, neither the isomorphism theorem nor its proof give a hint as to how to find such finite sets of rationals. But if we note that the existence of the solution ϕ is provable in ERNA, as shown in Section 3, then we can use the construction applied in the consistency proof, as described in Section 4, to get a finite model that satisfies at least some of the desired properties. In particular, the finite model obtained through the consistency proof construction will be a solution in the sense of (2). Furthermore, if relations of the forms (3) are provable in ERNA, we can carry out the consistency proof construction with the assurance that they will hold in our finite model. Through this we can see, not only the very constructive nature of ERNA, but also the way in which ERNA allows us to do many specific mathematical applications without using properties special to the continuum.

REFERENCES

Albeverio, S. *et al.* (1986). *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*. New York. Academic Press.

Chuaqui, R., and Suppes P., (1995). Free-variable axiomatic foundations of infinitesimal analysis. A fragment with finitary consistency proof. *Journal of Symbolic Logic*, 60 (1), 122–159.

Coddington, E. A., & Levinson, N., (1995). *Theory of Ordinary Differential Equations*. New York. McGraw Hill.

Hurd A. E. and Loeb, P. A., (1985). *An Introduction to Nonstandard Real Analysis*. New York. Academic Press.

Rose, H. E., (1984). *Subrecursion Functions and Hierarchies*. Oxford. Clarendon Press.

Simpson, S., (1984). Which set existence axioms are needed to prove the Cauchy/Peano theorem for ordinary differential equations? *Journal of Symbolic Logic*, 49 (3), 783–802.

Sommer, R., and Suppes, P., (1996). Finite models of elementary recursive nonstandard analysis. *Notas de la Sociedad Matemática de Chile*, 15(1).

Suppes, P., and Chuaqui, R., (1993). A finitarily consistent free-variable positive fragment of infinitesimal analysis, in *Proceedings, IX Latin American Symposium on Mathematical Logic, Universidad Nacional del Sur, Bahia Blanca, Argentina*, 1993. Notas de Logica Mathematica, Vol. 38, pp. 1–59.

Takeuti, G., (1987). *Proof Theory*. Amsterdam. North Holland.

Received November 8, 1996

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium