



NONSTANDARD METHODS AND DISCOVERY IN NUMBER THEORY AND COMBINATORICS

David Ross

Department of Mathematics, University of Hawai'i, Honolulu, Hawai'i,
 ross@math.hawaii.edu

Received: 9/8/25, Accepted: 5/8/26, Published: 9/28/26

Abstract

Many people still think of nonstandard analysis primarily as a way to replace epsilon-delta arguments by a rigorous use of infinitesimals. However, there are other very powerful properties this machinery can bring to arguments, and even if a final proof is entirely standard it can be motivated by the nonstandard argument. We will describe some examples from number theory and combinatorics where a nonstandard argument suggests a standard proof, or otherwise gives an unexpected insight into the problem.

1. Introduction

Infinitesimals and infinite numbers have been used in number theory for centuries, even after they were declared *numeri non grati* by critics such as Bishop Berkeley. For example, Euler famously discovered the identity $\zeta(2) = \pi^2/6$ by treating $\sin(x)$ as an infinite polynomial (given by its Taylor series), factoring it using its roots, and equating coefficients.¹ The self-evident power of such arguments has inspired many mathematicians over the years to make infinitesimal reasoning rigorous, and while the early efforts of mathematicians such as du Bois-Reymond, Hadamard, Borel, Hardy et al. were only moderately successful (see Hardy [17]), Abraham Robinson's introduction of *Nonstandard Analysis* in 1961 gave a complete and rigorous framework for arguments such as these [42].

Most recently, the most striking applications of nonstandard methods in this field have been results in additive number theory and Ramsey theory (for example [23] and [10]). These results generally leverage a remarkable property of nonstandard measure theory, automatic σ -additivity of certain measures constructed nonstandardly, and rely on the nonstandard technology in an essential way.

DOI: 10.5281/zenodo.23018342

¹For a historical discussion of this derivation, and how it looks when viewed through the lens of nonstandard analysis, see the excellent paper [32] by Mark McKinzie and Curt Tuckey.

In these pages we want to return to the Eulerian idea of using infinite numbers and infinitesimals as a tool for exploration and discovery in number theory and combinatorics, where a nonstandard argument finally motivates a standard proof. Because of the occasion of this conference, we will focus on examples where Mel Nathanson and the author have had some discussion or collaboration.

To motivate what follows, consider the Diophantine equation:

$$a = \frac{a_1}{x_1} + \cdots + \frac{a_n}{x_n}.$$

This equation has an interesting history. O. D. Kellogg [25] (in 1921) and D. R. Curtiss [7] (in 1922) considered the case where $a = a_1 = \cdots = a_n = 1$. Thirty years later W. Sierpinski [50] considered the more general case of $a, a_1, \dots, a_n$ being arbitrary positive real numbers.

The following result is due to Sierpinski.

Theorem 1. *If $0 < a, a_1, \dots, a_n \in \mathbb{R}$ then this equation has only finitely many solutions in $\mathbb{N}$.*

Here is a simple proof using nonstandard analysis. We will say more about what that means later; for now, channel Euler and assume that $\mathbb{R}$ has infinite integers and infinitesimals, whatever those are.

So suppose the theorem is false; that is, there are infinitely many solutions. Then there is a solution with at least one nonstandard (so infinite) x_i . (This uses a magic property of nonstandard models called *overspill*; see the next section.) By reordering if necessary assume x_i is finite for $i \leq r$, x_i infinite otherwise; so:

$$a = \underbrace{\frac{a_1}{x_1} + \frac{a_2}{x_2} + \cdots + \frac{a_r}{x_r}}_{\text{standard (possibly 0)}} + \underbrace{\frac{a_{r+1}}{x_{r+1}} + \cdots + \frac{a_n}{x_n}}_{\text{positive infinitesimal}}.$$

But this is impossible: the sum of the r standard numbers is standard, the sum of the $n-r$ positive infinitesimals is a positive infinitesimal, and two standard numbers cannot differ by a positive infinitesimal.²

2. What is Nonstandard Analysis?

The readers of this paper will be familiar with the idea of extending a mathematical system in order to facilitate understanding of the original system, for example the

²This proof is not mine, and the author does not know to whom it should be credited. The author has used it as an example on several occasions, including at a workshop at the American Institute of Mathematics in 2017. Mel Nathanson was at that workshop, and afterwards we had several discussions about this and related results of Sierpinski; this led to extensions by Nathanson [34] and by the author [48, 49].

extension of $\mathbb{R}$ to $\mathbb{C}$. Nonstandard analysis is a toolkit borrowed from mathematical logic to do this very generally. While it was established in approximately its current form by Abraham Robinson in 1961, it is built on ideas that go back at least to Thoralf Skolem's [51] elementary extensions of $\mathbb{N}$ in 1934.

The reader is referred to any one of many excellent introductions to nonstandard analysis, for example the comprehensive text [3], the obsolete but very readable introduction by Martin Davis [9], the brief introduction [27] by Tom Lindstrøm in [8], or even Robinson's original book [42]. In this section we will give a very superficial introduction to this toolkit, hopefully enough so that readers can follow the applications in the sections that follow.

Start with a *standard universe* (**SU**), which is a collection of all the mathematical objects we want to extend. For example, the **SU** will contain numbers $(0, 1, \dots; \sqrt{2}, e, \pi, \dots)$; it will also contain sets of numbers (such as $\mathbb{N}$ and $\mathbb{R}$) as *objects*; functions $(\sin : \mathbb{R} \rightarrow \mathbb{R}, + : \mathbb{R}^2 \rightarrow \mathbb{R}, \dots)$, function spaces $(\ell^p, \mathcal{C}(\ell^q \times \ell^p, \mathbb{R}))$ and so on.³

We now posit the existence of a *nonstandard universe* (**NSU**) extending the **SU**, and a mapping $*$: **SU** $\rightarrow$ **NSU**, satisfying some properties. (In other words, the **NSU** will contain objects such as $^*\mathbb{N}$, $^*\mathbb{R}$, $^*\sin$, $^*\ell^p$, $^*\mathcal{C}(\ell^q \times \ell^p, \mathbb{R})$, and so on. It will also contain *0 , *1 , $^*\sqrt{2}$, *e , $^*\pi$, $^*+$, etc., but we usually drop the stars on simple objects such as these.)

What might be called the *Fundamental Theorem of Nonstandard Analysis* is now the following.

Theorem 2. *Let **SU** be a standard mathematical universe. There exists a non-standard mathematical universe **NSU**, and a function $*$: **SU** $\rightarrow$ **NSU** satisfying:*

I. Transfer. *Every first-order statement P is true in the **SU** if and only if its 'star' is true in the **NSU**.*

II. Saturation. *Suppose $\{A_i\}_{i \in I}$ is a family of sets in the **SU**, indexed by an index set in the **SU**. Suppose moreover that $\{A_i\}_{i \in I}$ satisfies the **finite intersection property**:*

$$\text{For every finite nonempty } I' \subseteq I, \quad \bigcap_{i \in I'} A_i \neq \emptyset.$$

Then

$$\bigcap_{i \in I} ^*A_i \neq \emptyset.$$

The Transfer Property says intuitively that the **SU** and **NSU** have the 'same truths'. (Logicians would say that the nonstandard universe is an *elementary extension* of the standard universe.) *First-order* statements are those that can be

³One way of *constructing* a **SU** is to start with a model of set theory, and then interpret all the objects of mathematics in it in the usual foundational way. For our purposes it is easier to simply assume that we have such a universe, that it satisfies obvious closure properties and includes all the mathematical objects we will need.

expressed as a finite first-order sentence in a language whose symbols are elements of the **SU**, as well as a distinguished symbol ϵ . The star of a first-order statement is the one obtained by replacing every such symbol by its embedding under the $*$ function. We will see some examples below.

Saturation says intuitively that anything that *can* happen in the **SU**, *does* happen in the **NSU**. As expressed here this is an extremely weak form of saturation (often called an *enlargement*), but it is all we need for this paper.

Here are some simple examples to illustrate Transfer and Saturation.

1. $3 \in \mathbb{R}$, so $*3 \in *\mathbb{R}$. (Though as noted above we usually drop the $*$ from simple objects like 3, so we would just write $3 \in *\mathbb{R}$.)
2. $(\forall x \in \mathbb{N})[x \in \mathbb{R}]$, so $(\forall x \in *\mathbb{N})[x \in *\mathbb{R}]$. It follows that $*\mathbb{N} \subseteq *\mathbb{R}$.
3. $(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})[x + y = y + x]$, so $(\forall x \in *\mathbb{R})(\forall y \in *\mathbb{R})[x^* + y = y^* + x]$.
4. $*\mathbb{R}$ is an ordered field under the stars of the usual operations. (Note that we usually drop the $*$ from the operations on $*\mathbb{R}$.)
5. $*\mathbb{R}$ contains positive infinitesimals (by Saturation): $\bigcap_{n \in \mathbb{N}^+} *(0, 1/n) \neq \emptyset$.
6. $*$ preserves finite Boolean operations. For example, suppose $C = A \cap B$. Then

$$(\forall x \in C)[x \in A \wedge x \in B] \bigwedge (\forall x \in A)[x \in B \Rightarrow x \in C],$$

so

$$(\forall x \in *C)[x \in *A \wedge x \in *B] \bigwedge (\forall x \in *A)[x \in *B \Rightarrow x \in *C];$$

that is, $*C = *A \cap *B$.

7. If A is a finite set in the **SU**, say $A = \{a_1, \dots, a_n\}$, then

$$(\forall x \in A)[x = a_1 \vee x = a_2 \vee \dots \vee x = a_n], \text{ so } (\forall x \in *A)[x = *a_1 \vee x = *a_2 \vee \dots \vee x = *a_n].$$

In other words, finite sets do not get bigger in the nonstandard universe. (In fact, $*A = \{^*a_1, \dots, ^*a_n\}$.)

8. On the other hand, if A is an infinite set in the **SU** then

$$(*A) \setminus \{^*a : a \in A\} = \bigcap_{a \in A} (*A) \setminus \{^*a\} = \bigcap_{a \in A} *(A \setminus \{a\}) \neq \emptyset$$

by Saturation; in other words, infinite sets do get bigger in the nonstandard universe. For example, $*\mathbb{N}$ contains all the standard (or *finite*) natural numbers, but it also contains some new elements. If H is such a new element, it must appear in the $*$ -ordering after all the finite numbers. To see why this is true, consider any finite natural number, say 17 for definiteness. Then

$$(\forall x \in \mathbb{N})[x < 17 \Rightarrow (x = 0 \vee x = 1 \vee \dots \vee x = 16)],$$

so by Transfer and the fact that $H \neq 0, 1, \dots, 16, 17$, $H > 17$. These new elements of $*\mathbb{N}$ are therefore usually called *infinite* natural numbers. We often refer to an element of $*\mathbb{N}$ as a *hyperfinite number*.

9. The *Well-Ordering Property* of the natural numbers is often formulated as:

$$(\forall A \subseteq \mathbb{N})[A \neq \emptyset \Rightarrow (\exists x \in A)[(\forall y \in A)[x \leq y]]].$$

However, this is not a first-order formula (because of the quantification over subsets, $(\forall A \subseteq \mathbb{N})$), so does not transfer. The first-order formula

$$(\forall A \in \mathcal{P}(\mathbb{N}))[A \neq \emptyset \Rightarrow (\exists x \in A)[(\forall y \in A)[x \leq y]]]$$

does transfer, and so every nonempty element of ${}^*\mathcal{P}(\mathbb{N})$ has a least element.⁴

The reader will note that all the first-order formulas in these examples have *bounded quantifiers* (such as the $(\forall x \in C)$ in Example 6); for technical reasons we include this restriction in the definition of *first-order formula*.

Recognizing which mathematical statements are indeed first-order, and therefore which subsets of a set in the **NSU** are themselves in the **NSU**, is a skill analogous to recognizing which functions are measurable in measure theory, or which sets are Baire in general topology.

When applying nonstandard analysis it is sometimes easier to work with consequences of Saturation, rather than with the full principle. For example:

Overflow/Overspill: A first-order property true for arbitrarily large finite numbers is true for arbitrarily small infinite numbers.

Underspill/Underflow: A first-order property true for arbitrarily small infinite numbers is true for arbitrarily large finite numbers.

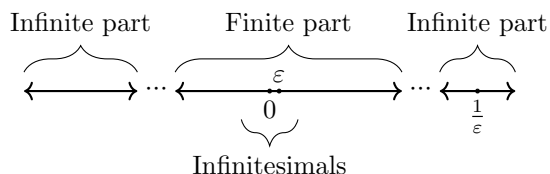
These are actually consequences of saturation: consider $\bigcap_{n \in \mathbb{N}} {}^*(n, \infty)$. These principles say that the line between finite and infinite numbers is blurry. We have already used overspill in the example at the end of Section 1.

2.1. The Structure of ${}^*\mathbb{R}$

By Transfer ${}^*\mathbb{R}$ is an ordered field, and as noted above it contains infinitesimals and (since it is a superset of ${}^*\mathbb{N}$) infinite elements. As above, an *infinitesimal* is an element of $\bigcap_{0 < n \in \mathbb{N}} {}^*(-1/n, 1/n)$. Equivalently, x is infinitesimal if $|x| < 1/n$ for every finite positive natural number n . Of course, 0 is an infinitesimal under this definition, but Saturation assures the existence of nonzero infinitesimals as well. (In fact, any proper ordered field extension of $\mathbb{R}$ contains such numbers.) We can visualize ${}^*\mathbb{R}$ as follows.

⁴It follows that the set ${}^*\mathbb{N} \setminus \mathbb{N}$ of *infinite integers* is not in ${}^*\mathcal{P}(\mathbb{N})$, even though it is a subset of ${}^*\mathbb{N}$. This means it is not in the **NSU**, since one can show that every element of the nonstandard universe which is a subset of $\mathbb{N}$ is in ${}^*\mathcal{P}(\mathbb{N})$.

A useful principle called the *Internal Definition Principle*, which follows from Transfer, is that if A is in the **NSU**, and B is defined from A using first-order formulas, then B is in the **NSU**. For example, if A is a nonstandard group, then the centralizer of A is also in the **NSU**.



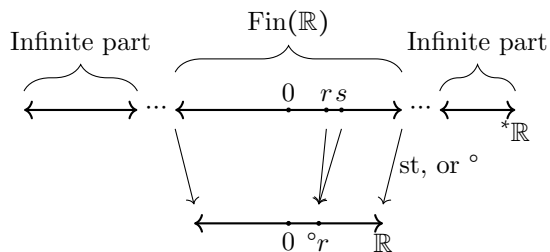
Write $\text{Fin}(\mathbb{R})$ for the finite part of ${}^*\mathbb{R}$:

$$\text{Fin}(\mathbb{R}) = \{x \in {}^*\mathbb{R} : |x| < n \text{ for some finite } n \in \mathbb{N}\}.$$

$\text{Fin}(\mathbb{R})$ is a subring of $\mathbb{R}$, but not a field: the reciprocal of a nonzero infinitesimal is infinite. The infinitesimals form a subring of $\text{Fin}(\mathbb{R})$ and so of course of ${}^*\mathbb{R}$, but are also an ideal in $\text{Fin}(\mathbb{R})$: if x is finite and ϵ is infinitesimal then ϵx is infinitesimal. Write $x \approx y$ if $x - y$ is infinitesimal. This is an equivalence relation (on all of ${}^*\mathbb{R}$, not just $\text{Fin}(\mathbb{R})$).

An easy argument, using metric completeness of $\mathbb{R}$, shows that every $x \in \text{Fin}(\mathbb{R})$ differs by an infinitesimal from a unique standard element of $\mathbb{R}$; write $\text{st}(x)$ or ${}^\circ x$ for that element, the *standard part* of x . We will use the two notations st and ${}^\circ$ interchangeably, as convenient, for this standard part map. Observe that st is the quotient map for the $\approx$ relation, it is a ring homomorphism from $\text{Fin}(\mathbb{R})$ to $\mathbb{R}$, ${}^\circ x \approx x$ for all $x \in \text{Fin}(\mathbb{R})$, and ${}^\circ x = x$ for all $x \in \mathbb{R}$.

In the following picture assume $r \approx s$.



2.2. Extending to ${}^*\mathbb{C}$ and Beyond

A similar picture to the above exists for the nonstandard extension ${}^*\mathbb{C}$ of the complex numbers. By transfer ${}^*\mathbb{C}$ is algebraically closed, it contains both $\mathbb{C}$ and ${}^*\mathbb{R}$ as subfields, it contains both infinitesimal and infinite elements (those with infinitesimal, respectively infinite, norms), there is a standard part map (which we again denote by st or ${}^\circ$) which is a ring homomorphism from the finite elements of ${}^*\mathbb{C}$ to $\mathbb{C}$.

More generally, suppose X is a topological space, with topology $\mathfrak{T}$. If $x \in X$ is standard, write:

$$\text{monad}(x) = \bigcap \{ {}^*u : x \in u, u \in \mathfrak{T} \}, \text{ the monad of } x.$$

In the case of $\mathbb{R}$ or $\mathbb{C}$, the monad of x is the set of elements of ${}^*\mathbb{R}$ or ${}^*\mathbb{C}$ differing from x by an infinitesimal.

An element y of *X is *nearstandard* if $y \in \text{monad}(x)$ for some $x \in X$. Denote by $NS(X)$ the nearstandard elements of *X , $NS(X) = \bigcup_{x \in X} \text{monad}(x)$. This is the analogue of $\text{Fin}(\mathbb{R})$ or $\text{Fin}(\mathbb{C})$.

The following are fundamental properties of nonstandard topology. The reader is referred to previously-referenced texts for the proof.

Lemma 1. *Let X be a topological space with topology $\mathfrak{T}$. Then:*

1. $\mathfrak{T}$ is Hausdorff if and only if for every $x \neq y$ in X , $\text{monad}(x) \cap \text{monad}(y) = \emptyset$.
2. $U \subseteq X$ is open if and only if for every $x \in U$, $\text{monad}(x) \subseteq {}^*U$.
3. (Robinson Compactness Criterion) $K \subseteq X$ is compact if and only if every element of *K is nearstandard; that is, ${}^*K \subseteq NS(X)$.

In particular, if $\mathfrak{T}$ is a Hausdorff topology on X then there is a well-defined standard part map $\text{st} : NS(X) \rightarrow X$, and $K \subseteq X$ is compact if and only ${}^*K \subseteq \text{st}^{-1} K$. This will be especially useful in Section 4.

3. Roots of Polynomials

The coefficients of a polynomial with complex coefficients depend continuously on its roots; that is an immediate consequence of the elementary *Viète Formulas*, known in some form as long ago as the 16th century.

The inverse, that the roots depend continuously on the coefficients, remains true, but is surprisingly tricky to prove—even to formulate correctly!—since, for example, a slight perturbation of the coefficients can cause roots to coalesce. Most of the proofs in the standard literature employ relatively heavy machinery (see, for example, [6, 18, 21, 22, 31, 39, 43, 53]).

At the conference Combinatorial and Additive Number Theory (CANT2022), Mel Nathanson asked:

Is there an easy proof that the roots of a polynomial depend continuously on the coefficients?

Recently the author, in collaboration with Nathanson, produced an entirely elementary proof [35] of this result. While that paper uses nothing beyond high school algebra, the current author's contribution was motivated by an argument using infinitesimals. This section is an exposition of the infinitesimal approach. While the

proof here carries less information than the standard proof, the lack of necessity for careful estimates make the proof shorter and arguably more transparent. (See Subsection 3.2 at the end of this section for a further discussion.)

3.1. Results

Let

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n$$

be a polynomial with complex coefficients and $a_n \neq 0$. By the Fundamental Theorem of Algebra, $f(z)$ fully factors as

$$f(z) = a_n \prod_{1 \leq i \leq n} (z - r_i)$$

where $r_1, r_2, \dots, r_n$ are the roots of $f(z)$, possibly with duplication.

Of course it suffices that any complex polynomial $f(z)$ has a single root in $\mathbb{C}$; the full factorization is then an induction on degree. For that reason, the fact that the extension ${}^*\mathbb{C}$ is algebraically closed shows that polynomials with coefficients in ${}^*\mathbb{C}$ can likewise be fully factored. Alternately, this is the subring ${}^*\mathbb{C}[x]$ of the ring ${}^*(\mathbb{C}[x])$ of *nonstandard* polynomials, elements of which can, by Transfer, be fully factored into a hyperfinite product. However, we will only consider the nonstandard polynomials of finite degree ${}^*\mathbb{C}[x]$.

In particular, suppose

$$g(z) = b_0 + b_1 z + \cdots + b_n z^n = b_n \prod_{1 \leq i \leq n} (z - s_i)$$

is a nonstandard polynomial, so each b_i is in ${}^*\mathbb{C}$. Say that $g(z)$ is an *infinitesimal deformation* of $f(z)$ provided $a_i \approx b_i$ for $0 \leq i \leq n$. Say that $g(z)$ is *infinitesimally aligned* with $f(z)$ provided there are orderings $r_1, \dots, r_n$ and $s_1, \dots, s_n$ of the (not necessarily distinct) roots of $f(z)$ and $g(z)$, respectively, such that $s_i \approx r_i$ for $1 \leq i \leq n$. The main result of this section is the following.

Theorem 3. *Suppose*

$$g(z) = b_0 + b_1 z + \cdots + b_n z^n = b_n \prod_{1 \leq i \leq n} (z - s_i)$$

is an infinitesimal deformation of

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n = a_n \prod_{1 \leq i \leq n} (z - r_i)$$

(where $a_n \neq 0$). Then $g(z)$ is infinitesimally aligned with $f(z)$.

Note that the wording of Theorem 3 references nonstandard elements; this would have been a natural way of stating the result before the rise of epsilonics. It is a straightforward exercise in nonstandard analysis to show that this result is equivalent to the following standard one.

Theorem 4. *Suppose*

$$f(z) = a_0 + a_1z + \cdots + a_nz^n = a_n \prod_{1 \leq i \leq n} (z - r_i)$$

is a polynomial over $\mathbb{C}$, where $a_n \neq 0$. For every $\epsilon > 0$ there exists a $\delta > 0$ such that whenever

$$g(z) = b_0 + b_1z + \cdots + b_nz^n = b_n \prod_{1 \leq i \leq n} (z - s_i)$$

is polynomial with $|a_i - b_i| < \delta$ for $0 \leq i \leq n$, then there are orderings $\{r_i\}_{i=1}^n$ and $\{s_i\}_{i=1}^n$ of the roots of f and g (respectively), possibly with duplication, so that $|r_i - s_i| < \epsilon$ for all i .

(However, the proof of the equivalence does require properties of the nonstandard model such as Transfer and some form of Saturation; in the following proof of Theorem 3, all that are needed are some basic arithmetic and algebraic properties of the extension ${}^*\mathbb{C}$ and the standard part map.)

Proof 1 of Theorem 3. Suppose $z \in \mathbb{C}$ is standard. Then:

$${}^\circ g(z) = \left(\sum_{i=0}^n b_i z^i \right) = \sum_{i=0}^n {}^\circ b_i z^i = \sum_{i=0}^n a_i z^i = f(z) = a_n \prod_{1 \leq i \leq n} (z - r_i).$$

But also:

$${}^\circ g(z) = \left(b_n \prod_{1 \leq i \leq n} (z - s_i) \right) = {}^\circ b_n \prod_{1 \leq i \leq n} (z - {}^\circ s_i)$$

which, by the uniqueness of the factorization in the Fundamental Theorem of Algebra, seems to prove the result. Well, not quite; we need to show that the roots of g are finite, so ${}^\circ s_i$ exists.

So, suppose s is an infinite root of g . Then:

$$0 = \frac{g(s)}{s^n} = \underbrace{b_n}_{\text{not infinitesimal}} + \underbrace{\frac{b_{n-1}}{s} + \cdots + \frac{b_0}{s^n}}_{\text{infinitesimal}}$$

which of course cannot happen. $\square$

As mentioned, it is not hard to show that the nonstandard statement Theorem 3 is equivalent to the standard Theorem 4, but of course then you are still stuck with a nonstandard proof. What if we are interested (as was the original goal) in an elementary standard proof? Unfortunately, while the above argument is apparently quite simple, it is very nonconstructive, and does not have an obvious standard counterpart.

In order to extract a standard proof from these ideas, we start with a nonstandard proof of Theorem 3 which is more involved than the one above, but also more constructive; it approximates the roots one at a time, instead of all at once.

We begin with two lemmas.

Lemma 2. Suppose $f(z) = \sum_{i=0}^n a_i z^i$, where $a_n \neq 0$, is a polynomial with coefficients in $\mathbb{C}$, and $g(z) = \sum_{i=0}^n b_i z^i$ is a polynomial with coefficients in ${}^*\mathbb{C}$. Consider the statements:

- (i) $g(z)$ is infinitesimally aligned with $f(z)$ and $b_n \approx a_n$.
- (ii) $g(z) \approx f(z)$ for all $z \in \mathbb{C}$.
- (iii) $g(z)$ is an infinitesimal deformation of $f(z)$.

Then $(i) \Rightarrow (ii) \Leftrightarrow (iii)$.

Proof. Assume (i) holds. For any (standard) $z \in \mathbb{C}$:

$$\text{st } g(z) = \text{st} \left(b_n \prod_{1 \leq i \leq n} (z - s_i) \right) = \text{st } b_n \prod_{1 \leq i \leq n} (z - \text{st } s_i) = a_n \prod_{1 \leq i \leq n} (z - r_i) = f(z).$$

This proves $(i) \Rightarrow (ii)$.

Now, suppose (ii) holds. Suppose first that we know that all the coefficients of $g(z)$ are finite. For any (standard) $z \in \mathbb{C}$,

$$\sum_{i=0}^n \text{st } b_i z^i = \text{st} \left(\sum_{i=0}^n b_i z^i \right) = \text{st } g(z) = f(z) = \sum_{i=0}^n a_i z^i$$

and two polynomials agree on $\mathbb{C}$ only if they have the same coefficients; that is, $b_i \approx a_i$ for every i , proving $(ii) \Rightarrow (iii)$. It remains to show that if $g(z)$ takes only finite values for standard z , then it has finite coefficients. If not, then there is a coefficient $j \leq n$ with $|b_j|$ largest. But then for any (standard) $z \in \mathbb{C}$,

$$0 = \text{st} \left(\frac{g(z)}{b_j} \right) = \text{st} \left(\sum_{i=0}^n \frac{b_i}{b_j} z^i \right) = \sum_{i=0}^n \text{st} \left(\frac{b_i}{b_j} \right) z^i$$

so $\text{st}(b_i/b_j)$ must be 0 for all i ; but $\text{st}(b_j/b_j) = 1$, a contradiction.

Finally, suppose that (iii) holds, and $z \in \mathbb{C}$ is standard. Then

$$\text{st } g(z) = \text{st} \left(\sum_{i=0}^n b_i z^i \right) = \sum_{i=0}^n \text{st } b_i z^i = \sum_{i=0}^n a_i z^i = f(z)$$

proving $(iii) \Rightarrow (ii)$. □

Remark 1. (1) In the proof of $(ii) \Rightarrow (iii)$, it is only required that $g(z) \approx f(z)$ for infinitely many values (or even at least n different values) of z . (2) The implication $(i) \Rightarrow (iii)$ in this lemma is also a consequence of the Viète formulas, which give the coefficients of the polynomial as elementary symmetric functions of the roots. (3) The requirement that $b_n \approx a_n$ in (i) is necessary to account for the fact that the roots of a polynomial only determine the coefficients up to a constant multiple.

The second lemma is the key to the full result.

Lemma 3. Suppose $g(z) = \sum_{i=0}^n b_i z^i$ is an infinitesimal deformation of $f(z) = \sum_{i=0}^n a_i z^i$, where $a_n \neq 0$. Then for every root s of $g(z)$, $r = \text{st } s$ is a root of $f(z)$.

Proof. First, we note that s is not infinite; see the first proof (above) of this theorem. Therefore, s has a standard part $\text{st } s$, and

$$f(\text{st } s) = \sum_{i=0}^n a_i (\text{st } s)^i = \sum_{i=0}^n \text{st } b_i (\text{st } s)^i = \text{st} \left(\sum_{i=0}^n b_i s^i \right) = \text{st } 0 = 0.$$

So $r = \text{st } s$ is a root of $f(z)$. □

Proof 2 of Theorem 3. The proof is by induction on n . We are given that

$$g(z) = b_0 + b_1 z + \cdots + b_n z^n = b_n \prod_{1 \leq i \leq n} (z - s_i)$$

is an infinitesimal deformation of

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n = a_n \prod_{1 \leq i \leq n} (z - r_i).$$

When $n = 1$ the only root of $f(z)$ is $r_1 = -a_0/a_1$, and the only root of $g(z)$ is $s_1 = -b_0/b_1$. Moreover, $\text{st}(-b_0/b_1) = -\text{st } b_0/\text{st } b_1 = -a_0/a_1$ by the usual rules of nonstandard arithmetic on ${}^*\mathbb{C}$. It follows that $s_1 \approx r_1$.

Now, suppose the result is true for polynomials of degree $n - 1$. Let s be a root of $g(z)$. Since $g(z)$ is an infinitesimal deformation of $f(z)$, by Lemma 3 there is a root r of $f(z)$ with $r \approx s$. Write $f(z) = (z - r)\hat{f}(z)$ and $g(z) = (z - s)\hat{g}(z)$, where $\hat{f}$ and $\hat{g}$ are polynomials of degree $n - 1$ over their respective fields. Then for any $z \in \mathbb{C}$:

$$(z - r)(\hat{f}(z) - \hat{g}(z)) = f(z) - g(z) - \hat{g}(z)(s - r). \quad (1)$$

When $z \neq r$, $\hat{g}(z) = g(z)/(z - s)$ is finite, so the right hand side of (1) is infinitesimal, so $\hat{f}(z) - \hat{g}(z) \approx 0$, or $\hat{f}(z) \approx \hat{g}(z)$. By Lemma 2 (and the remarks following) $\hat{g}(z)$ is an infinitesimal deformation of $\hat{f}(z)$.

By the induction hypothesis there are orderings $r_1, r_2, \dots, r_{n-1}$ and $s_1, s_2, \dots, s_{n-1}$ of the roots of $\hat{f}(z)$, respectively $\hat{g}(z)$, with $s_i \approx r_i$ for $1 \leq i \leq n - 1$. Put $r_n = r$ and $s_n = s$, then $s_i \approx r_i$ for $1 \leq i \leq n$, proving that $g(z)$ is infinitesimally aligned with $f(z)$. □

3.2. Remarks

Lemma 3 in this paper is the nonstandard analogue of Nathanson and Ross [35, Theorem 4], and plays a similar role in the later induction. The difference is subtle: in [35, Theorem 4], given a root of $f(z)$, we show that small perturbations $g(z)$ have nearby roots. Lemma 3, on the other hand, shows that given a root of a perturbation $g(z)$ of $f(z)$, there is a nearby root of $f(z)$. The proof in [35] has an

additional benefit: given a tolerance ϵ , it gives a constructive bound $\delta = \delta(\epsilon)$ so that if the perturbation is kept smaller than δ then its roots are within the tolerance ϵ of the roots of f . The author does not see any way to extract this bound from the proof in this paper.

Moreover, while the proofs in [35] apply to any algebraically closed field with an absolute value, the proofs here (especially that of Lemma 3) use the standard part map in an essential way; this requires the absolute value on the underlying field to be complete and Archimedean, so by Ostrowski's Theorem we must be working on $\mathbb{C}$.

In the second proof of Theorem 3, it was not necessary to give an explicit form for $\hat{f}(z)$ and $\hat{g}(z)$ (as we did in the corresponding section of [35]). Again, while this simplification is a byproduct of the infinitesimal approach, it loses information about the nature of the approximation that is present in the standard paper.

4. Finitely Many Implies Infinitely Many

Recently Mel Nathanson became interested in some results of Alexander Abian⁵ of the form:

If some property P is true for every finite subset of a set of objects, then P is true simultaneously for all the objects.

For example, in Nathanson [37] he examined a result of Abian and Eslami [2] about certain sets $\mathfrak{F}$ of equations in infinitely many variables, where it is shown that if every finite subset of $\mathfrak{F}$ has a common solution of some form then $\mathfrak{F}$ has a common solution of that form.

Nathanson coined a term for results of this form: *Finitely many implies infinitely many*.

Results of this form are familiar to mathematical logicians as *compactness* results. In this section we give nonstandard proofs for some of the results in [37], as well as others. In nonstandard analysis the compactness principle is built into the construction of the nonstandard model, and so these results become completely transparent with the nonstandard machinery. This also makes it possible to identify natural generalizations of the theorems; see, for example, Theorem 9 below.

Anyone familiar with nonstandard analysis cannot fail to note the similarity of these nonstandard proofs to Robinson's nonstandard proof of the Tychonoff Theorem. This provides a roadmap to using Tychonoff's Theorem to give standard proofs for not only the Abian–Eslami result, but also strong generalizations. This is a project we carry out in Nathanson and Ross [36].

⁵Abian is not best known for his mathematics, but rather for his proposal that we blow up the moon; see for example coverage in the *Des Moines Register* [11] and the *Wall Street Journal* [52].

Questions about infinite systems of equations in infinitely many variables go back to the early days of functional analysis; arguably the starting point for the whole field. The Abian–Eslami result is a relative of a 1913 result of F. Riesz [41]. Riesz’s proof, reproduced in Banach’s 1932 Theory of Linear Operations [5]⁶ uses something today we would recognize as a special case of the Hahn–Banach Extension Theorem (Hahn [13], Banach [4]), which should probably be attributed to Helly ([19]; see Narici and Beckenstein [33] for the historical discussion).

It should also be noted that nonstandard analysis has proved a useful tool in understanding the Hahn–Banach Theorem. In 1962 Wim Luxemburg [29] showed that the Boolean Prime Ideal Theorem implies Hahn–Banach in Zermelo–Fraenkel set theory without the Axiom of Choice. A few years later he refined the result [30] to show that Hahn–Banach is equivalent to the statement that every Boolean algebra has a $[0, 1]$ -valued measure. Both these proofs used nonstandard analysis. See David Pincus [40] for the proof that the Boolean Prime Ideal Theorem is strictly stronger than Hahn–Banach, as well as a comprehensive discussion of Luxemburg’s results and other equivalents and inequivalents to the Hahn–Banach Theorem.

4.1. Finite Polynomials

Theorem 5 (Abian [1]). *Let R be a finite ring, $\mathfrak{F}$ a set of (finite) polynomials over (possibly infinitely many) variables $\{x_i\}_{i \in I}$, and suppose every finite subset of $\mathfrak{F}$ has a common root in R :*

$$\forall p_1, \dots, p_n \in \mathfrak{F} \quad \exists \mathbf{x} = \{x_i\}_{i \in I} \subseteq R \text{ such that } p_k(\mathbf{x}) = 0 \quad (1 \leq k \leq n).$$

Then there exists $\mathbf{x} \in R^I$ such that $p(\mathbf{x}) = 0$ for every $p \in \mathfrak{F}$.

Proof of Theorem 5. By saturation there exists a $\mathbf{x} \in {}^*R^I$ such that ${}^*p(\mathbf{x}) = 0$ for every $p \in \mathfrak{F}$. Since R is finite, ${}^*R = R$. It follows that for $i \in I$ standard, $x_i \in R$, so for any standard p , $p(\mathbf{x}|_I) = {}^*p(\mathbf{x})$. In particular, $p(\mathbf{x}|_I) = 0$ for every $p \in \mathfrak{F}$. $\square$

This result does not easily extend to infinite rings. For example, Abian gives the following example:

$$\begin{array}{c} (x-1) - y_1^2 \\ (x-2) - y_2^2 \\ (x-3) - y_3^2 \\ \vdots \\ (x-n) - y_n^2 \\ \vdots \end{array}$$

⁶An extension of Banach’s 1922 thesis.

Any finite subset of these has a solution in $\mathbb{R}$, but the full set does not. What makes the example work is that the finite satisfiability of the equations requires arbitrarily large solutions. If we bound the solutions, then we do get a positive result.

Theorem 6. *Let $\mathfrak{F}$ a set of (finite) real polynomials over (possibly infinitely many) variables $\{x_i\}_{i \in I}$, and suppose every finite subset of $\mathfrak{F}$ has a common root in $[-M, M]^I$ for some M :*

$$\forall p_1, \dots, p_n \in \mathfrak{F} \exists \mathbf{x} = \{x_i\}_{i \in I} \subseteq [-M, M] \text{ such that } p_k(\mathbf{x}) = 0 \quad (1 \leq k \leq n).$$

Then there exists $\mathbf{x} \in [-M, M]^I$ such that $p(\mathbf{x}) = 0$ for every $p \in \mathfrak{F}$.

Proof of Theorem 6. By saturation there exists a $\mathbf{x} \in {}^*([-M, M]^I) \subseteq {}^*([-M, M])^{*I}$ such that ${}^*p(\mathbf{x}) = 0$ for every $p \in \mathfrak{F}$. For $i \in I$ standard, $|x_i| \leq M$, so $y_i = \text{st } x_i$ exists. Put $\mathbf{y} = \{y_i\}_{i \in I}$. Now, if $p \in \mathfrak{F}$ is standard then its coefficients are all standard and its variables all have standard indices; it follows that $p(\mathbf{y}) = \text{st } {}^*p(\mathbf{x}) = 0$, proving the theorem. $\square$

Remark 2. With the obvious modification Theorem 6 is true in $\mathbb{C}$ as well, but in fact Abian has showed that for $\mathbb{C}$ the bound is not necessary.

The proof works *mutatis mutandis* if we choose a different M_i for each i ; that is, change the hypothesis/conclusion to $\mathbf{x} \in \prod_{i \in I} [-M_i, M_i]$. In fact, this result has an obvious generalization.

Theorem 7. *Let $\mathfrak{F}$ a set of functions over (possibly infinitely many) variables $\{x_i\}_{i \in I}$ with the property that each $f \in \mathfrak{F}$ (a) only depends on finitely many variables in I , and (b) is continuous in each variable. Let $\mathcal{K}$ be a subset of $\mathbb{R}^I$ which is compact in the product topology. Suppose every finite subset of $\mathfrak{F}$ has a common root in $\mathcal{K}$:*

$$\forall f_1, \dots, f_n \in \mathfrak{F} \exists \mathbf{x} = \{x_i\}_{i \in I} \in \mathcal{K} \text{ such that } f_k(\mathbf{x}) = 0 \quad (1 \leq k \leq n).$$

Then there exists $\mathbf{x} \in \mathcal{K}$ such that $p(\mathbf{x}) = 0$ for every $p \in \mathfrak{F}$.

4.2. Functions of Infinitely Many Variables

What if the functions are functions of infinitely many variables? The functional analysts of the early 20th century understood that even for linear functions of infinitely many variables, constraints needed to be added.

For example, Helly [20] offered the following cautionary example in 1921:

$$\begin{aligned} 1 &= x_1 + x_2 + x_3 + x_4 + x_5 + \cdots \\ 1 &= \quad x_2 + x_3 + x_4 + x_5 + \cdots \\ 1 &= \quad \quad x_3 + x_4 + x_5 + \cdots \\ &\vdots \end{aligned}$$

We now give a short nonstandard proof of the following theorem of Abian and Eslami [2].

Theorem 8 (Abian and Eslami 1982). *Let $M > 0$, and $p, q > 1$ be conjugate, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Let I be an infinite set, $\mathbf{a}_i \in \ell^p$ and $b_i \in \mathbb{R}$ for all $i \in I$, and suppose that for every finite $I' \subseteq I$ the equations*

$$\sum_{n=0}^{\infty} a_{in} x_n = b_i \quad (i \in I')$$

have a simultaneous solution $\mathbf{x} \in \ell^q$ with $\|\mathbf{x}\|_q \leq M$. (Equivalently, for every finite $I' \subseteq I$ there exists $\mathbf{x} \in \ell^q$ with $\|\mathbf{x}\|_q \leq M$ and $\mathbf{a}_i \cdot \mathbf{x} = b_i$ for $i \in I'$.) Then there exists $\mathbf{x} \in \ell^q$ with $\|\mathbf{x}\|_q \leq M$ and $\mathbf{a}_i \cdot \mathbf{x} = b_i$ for all $i \in I$.

Adopt the notation $\mathbf{a} = \{a_n\}_{n \in \mathbb{N}}$ for a sequence of real numbers, $\mathbf{a}_i = \{a_{in}\}_{n \in \mathbb{N}}$ for a sequence of sequences of real numbers indexed by i , and $\mathbf{a} \cdot \mathbf{b} = \sum_{n \in \mathbb{N}} a_n b_n$. If $N \in \mathbb{N}$ let $\mathbf{a}^N$ be the sequence:

$$a_n^N = \begin{cases} a_n & \text{if } n > N, \\ 0 & \text{otherwise.} \end{cases}$$

First, we need a few simple facts about ℓ^q (and of course ℓ^p).

Proposition 1. *Let $\mathbf{c} \in \ell^q$ and $N \in \mathbb{N}$. Then:*

- (i) $\mathbf{c}^N \in \ell^q$
- (ii) $\lim_{N \rightarrow \infty} \|\mathbf{c}^N\|_q = 0$
- (iii) $|c_N| \leq \|\mathbf{c}\|_q$

These are obvious.

Proof of Theorem 8. By Saturation there is an $\mathbf{x} \in {}^*\ell^q$ such that $\|x\|_q \leq M$ and for every (standard) $i \in I$, ${}^*\mathbf{a}_i \cdot \mathbf{x} = b_i$. Define a standard sequence y_n , $n \in \mathbb{N}$, by putting $y_n = \text{st } x_n$; this exists by Proposition 1(iii). It remains to prove three things:

- 1. $\mathbf{y} \in \ell^q$
- 2. $\|\mathbf{y}\|_q \leq M$
- 3. $\mathbf{a}_i \cdot \mathbf{y} = b_i$ for (standard) $i \in I$

For any $N \in \mathbb{N}$, $\sum_{n \leq N} |y_n|^q \approx \sum_{n \leq N} |x_n|^q \leq M^q$; letting $N \rightarrow \infty$ proves (1) and (2).

For (3), let N be a standard positive integer. Then

$$\begin{aligned}
 |b_i - \sum_{n \leq N} a_{in} y_n| &\approx |b_i - \sum_{n \leq N} a_{in} x_n| \quad (\text{since } N \text{ is finite and } y_n \approx x_n) \\
 &= \sum_{n > N, n \in {}^*\mathbb{N}} a_{in} x_n \quad (\text{since } {}^*\mathbf{a}_i \cdot \mathbf{x} = b_i) \\
 &= |{}^*\mathbf{a}_i^N \cdot \mathbf{x}| \quad (\text{by definition of } \mathbf{a}_i^N) \\
 &\leq \|{}^*\mathbf{a}_i^N\|_p \|\mathbf{x}\|_q \quad (\text{by } {}^*\text{H\"older}) \\
 &\leq \|\mathbf{a}_i^N\|_p M \\
 &\rightarrow 0 \text{ as } N \rightarrow \infty \quad (\text{by Proposition 1(ii)})
 \end{aligned}$$

which proves (3), and the theorem. $\square$

4.3. Generalizing the Result

We now consider two possible generalizations: easing the bound M (as discussed in the remarks after Theorem 6), and replacing the exact solution with an approximation. The following does both at once.

Theorem 9. *Let $p, q > 1$ be conjugate, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Let $M_n > 0$ for every $n \in \mathbb{N}$. Let I be an infinite set, $\mathbf{a}_i \in \ell^p$ and $b_i \in \mathbb{R}$ for all $i \in I$. Then the following are equivalent:*

1. *There exists an $\mathbf{x} \in \ell^q$ with $\mathbf{a}_i \cdot \mathbf{x} = b_i$ for all $i \in I$ and $|x_n| \leq M_n$ for all $n \in \mathbb{N}$.*
2. *There is a positive sequence $\{e_N\}_{N \in \mathbb{N}}$ with $\lim_{N \rightarrow \infty} e_N = 0$ such that for every finite $I' \subseteq I$ and every $\epsilon > 0$, there is an $\mathbf{x}$ with (a) $\|\mathbf{x}^N\|_q \leq e_N$ for all N , (b) $|x_n| \leq M_n$ for all $n \in \mathbb{N}$, and (c) $|\mathbf{a}_i \cdot \mathbf{x} - b_i| < \epsilon$ for $i \in I'$.*

Proof. (1 $\Rightarrow$ 2) Given $\mathbf{x}$, let $e_N = \|\mathbf{x}^N\|_q$. (2) is immediate.

(2 $\Rightarrow$ 1) By Saturation there is an $\mathbf{x} \in {}^*\ell^q$ such that (a) $\|\mathbf{x}^N\|_q \leq e_N$ for all standard N , (b) $|x_n| \leq M_n$ for all $n \in {}^*\mathbb{N}$ (including standard n), and (c) $|{}^*\mathbf{a}_i \cdot \mathbf{x} - b_i| \approx 0$ for $i \in I$.

Put $y_n = \text{st } x_n$ for standard n . As in the proof of Theorem 8, it remains to prove three things:

1. $\mathbf{y} \in \ell^q$
 2. $|y_n| \leq M_n$ for all $n \in \mathbb{N}$
 3. $\mathbf{a}_i \cdot \mathbf{y} = b_i$ for (standard) $i \in I$
- (2) is immediate. If $N < K$ are standard then

$$\sum_{n=N+1}^K |y_n|^q \approx \sum_{n=N+1}^K |x_n|^q \leq \|\mathbf{x}^N\|_q^q \leq e_N^q.$$

Let $K \rightarrow \infty$, obtain $\|\mathbf{y}^N\|_q \leq e_N$. By the earlier observation, $\mathbf{y} \in \ell^q$, proving (1).

For (3), let N be a standard positive integer, then

$$\begin{aligned} \left| b_i - \sum_{n \leq N} a_{in} y_n \right| &\leq |b_i - {}^* \mathbf{a}_i \cdot \mathbf{x}| + \left| \sum_{n \leq N} a_{in} x_n - \sum_{n \leq N} a_{in} y_n \right| + \left| {}^* \mathbf{a}_i \cdot \mathbf{x} - \sum_{n \leq N} a_{in} x_n \right| \\ &\approx |{}^* \mathbf{a}_i \cdot \mathbf{x}^N| \quad (\text{by definition of } \mathbf{x}^N) \\ &\leq \|{}^* \mathbf{a}_i\|_p \|\mathbf{x}^N\|_q \quad (\text{by } {}^* \text{H\"older}) \\ &\leq \|\mathbf{a}_i\|_p e_N \\ &\rightarrow 0 \text{ as } N \rightarrow \infty \end{aligned}$$

where for the second line the first term is infinitesimal since $x_n \approx y_n$ for every $n \leq N$ and the second term is infinitesimal by choice of $\mathbf{x}$. This proves (3), and the theorem. $\square$

Remark 3. In [36] Nathanson and the author were able to considerably strengthen this result, by replacing equations with approximate equations, and linear equations with certain nonlinear ones. The first generalization answered a question raised by a participant in the New York Number Theory Seminar, and the proof was suggested by the nonstandard proof.

5. Sidon and B_h -sets

Let G be a commutative semigroup (usually $\mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{N}, \dots$ under $+$) and $A \subseteq G$. If $h \in \mathbb{N}_+$,

$$hA = \underbrace{A + A + \dots + A}_{h \text{ copies}} := \{a_1 + \dots + a_h : c_i \in A\}.$$

The set A is a B_h -set in G provided every element of hA has only one representation (up to permutation of the terms of the sum). If $h = 2$ then this is a *Sidon set*.

Mel Nathanson recently observed the following.

Proposition 2 (Nathanson). *For $h \in \mathbb{N}_+$ let*

$$\chi_{h,n} = \{\mathbf{n} \in \mathbb{N}^n : \sum_i n_i = h\}.$$

Let $X = \{x_1, \dots, x_n\}$ be a finite subset of $\mathbb{R}$ and write $\mathbf{x} = \langle x_1, \dots, x_n \rangle \in \mathbb{R}_+^n$. Then X is a B_h set in $\mathbb{R}$ provided the values of

$$\mathbf{n} \cdot \mathbf{x} = \sum_i n_i x_i$$

are distinct for $\mathbf{n} \in \chi_{h,n}$.

In other words, for any n and h , there is a finite set of linear combinations with integer coefficients to avoid in order for a set A of size n to be a B_h -set. The following is immediate.

Corollary 1. *If $A \subseteq \mathbb{R}$ is linearly independent over $\mathbb{Q}$ then A is a B_h -set in $\mathbb{R}$.*

The following is harder.

Theorem 10 (Nathanson [38]). *For any finite $A \subseteq \mathbb{R}$ which is linearly independent over $\mathbb{Q}$, for large enough q the set of integer parts of $\{qa : a \in A\}$ forms a B_h -set over $\mathbb{Z}$.*

In this section we will give a simple nonstandard proof of this result, which turns out to be an immediate consequence of underspill; the hardest part is notation. We will also give a standard proof based on the nonstandard argument.

If $x \in \mathbb{R}_+$ and $H \in {}^*\mathbb{N}$, write

$$\lfloor x \rfloor_H = \frac{\lfloor 10^H x \rfloor}{10^H}.$$

If $\mathbf{x} = \langle x_1, \dots, x_n \rangle \in \mathbb{R}_+^n$ write $\lfloor \mathbf{x} \rfloor_H = \langle \lfloor x_1 \rfloor_H, \dots, \lfloor x_n \rfloor_H \rangle$.

We could use any positive integer (or even positive real number) in place of 10^H , but everything is easy to visualize with 10^H :

$$\begin{aligned} x &= d.d_1d_2d_3\dots d_Hd_{H+1}\dots \\ \lfloor x \rfloor_H &= d.d_1d_2d_3\dots d_H000\dots \end{aligned}$$

Proposition 3. (i) *If $x \in \mathbb{R}_+$ and H is standard then $10^H \lfloor x \rfloor_H$ is a positive integer.*

(ii) *If $x \in \mathbb{R}_+$ and H is infinite then $\lfloor x \rfloor_H \approx x$.*

(iii) *If $x_1, \dots, x_n \in \mathbb{R}_+$ are linearly independent over $\mathbb{Q}$ and $H \in {}^*\mathbb{N}$ is infinite then $\lfloor x_1 \rfloor_H, \lfloor x_2 \rfloor_H, \dots, \lfloor x_n \rfloor_H$ are linearly independent over $\mathbb{Q}$.*

(In (ii) and (iii) only H (and any subsequent hyperreals $\lfloor x_i \rfloor_H$) might be non-standard.)

Proof. Only (iii) is nontrivial. Suppose $\sum_{i=1}^n a_i \lfloor x_i \rfloor_H = 0$, $a_i \in \mathbb{Q}$. Then

$$0 = \text{st} \sum_{i=1}^n a_i \lfloor x_i \rfloor_H = \sum_{i=1}^n a_i \text{st} \lfloor x_i \rfloor_H = \sum_{i=1}^n a_i x_i$$

so every $a_i = 0$. □

Corollary 2 (Theorem 10, but for $q = 10^H$). *If $h \in \mathbb{N}_+$ and $x_1, \dots, x_n \in \mathbb{R}_+$ are linearly independent over $\mathbb{Q}$, then for some $H_0 \in \mathbb{N}$ and all $H \geq H_0$ the set of rational numbers $\{\lfloor x_1 \rfloor_H, \dots, \lfloor x_n \rfloor_H\}$ is a B_h -set in $\mathbb{R}$, and the set of positive integers $\{10^H \lfloor x_1 \rfloor_H, \dots, 10^H \lfloor x_n \rfloor_H\}$ is a B_h -set in $\mathbb{N}$.*

Proof. For any infinite H , $\mathbf{x} = \langle \lfloor x_1 \rfloor_H, \dots, \lfloor x_n \rfloor_H \rangle$ satisfies the first-order formula

$$\bigwedge_{\mathbf{m} \neq \mathbf{n} \in \chi_{h,n}} \mathbf{m} \cdot \mathbf{x} \neq \mathbf{n} \cdot \mathbf{x} \quad (2)$$

so by underflow it satisfies formula 2 for all large enough standard H . $\square$

It is not difficult to standardize this proof. Begin with a standard version of Proposition 3.

Proposition 4. (i) If $x \in \mathbb{R}$ and H is standard then $10^H \lfloor x \rfloor_H$ is an integer.

(ii) If $x \in \mathbb{R}$ then $\lim_{H \rightarrow \infty} \lfloor x \rfloor_H = x$.

Proof. Only (ii) is new:

$$|x - \lfloor x \rfloor_H| = \left| \frac{10^H x}{10^H} - \frac{\lfloor 10^H x \rfloor}{10^H} \right| \leq \frac{1}{10^H} \rightarrow 0 \text{ as } H \rightarrow \infty. \quad \square$$

Lemma 4. Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{c}^1, \dots, \mathbf{c}^m \in \mathbb{Q}^n$.

1. If $\mathbf{c}^j \cdot \mathbf{x} = \sum_{i=1}^n c_i^j x_i \neq 0$ for every $j \leq m$ then for some finite $H_0 \in \mathbb{N}$ and all $H \geq H_0$,

$$\sum_{i=1}^n c_i^j \lfloor x_i \rfloor_H \neq 0 \quad (j \leq m). \quad (3)$$

2. If the values of $\mathbf{c}^j \cdot \mathbf{x}$ are all distinct, $j \leq m$, then for some finite $H_0 \in \mathbb{N}$ and all $H \geq H_0$, the values of $\mathbf{c}^j \cdot \lfloor \mathbf{x} \rfloor_H$ are all distinct.

Proof of (1); (2) is similar. Otherwise there is an increasing sequence $\{H_k\}_{k \in \mathbb{N}}$, and a sequence $\{\mathbf{d}^k\}_{k \in \mathbb{N}} \subseteq \{\mathbf{c}^1, \dots, \mathbf{c}^m\}$ such that

$$\sum_{i=1}^n d_i^k \lfloor x_i \rfloor_{H_k} = 0.$$

Since $\{\mathbf{c}^1, \dots, \mathbf{c}^m\}$ is a finite set, at least one of these, say $\mathbf{c}^j$, must appear infinitely often in the sequence of $\mathbf{d}^k$. Thus,

$$\sum_{i=1}^n c_i^j \lfloor x_i \rfloor_{H_k} = 0$$

for infinitely many k . But for each $i \leq n$, $\lim_{k \rightarrow \infty} \lfloor x_i \rfloor_{H_k} = x_i$, so

$$\sum_{i=1}^n c_i^j x_i = 0,$$

a contradiction. $\square$

Corollary 3. If $h \in \mathbb{N}_+$ and $x_1, \dots, x_n \in \mathbb{R}_+$ are linearly independent over $\mathbb{Q}$, then for some $H_0 \in \mathbb{N}$ and all $H \geq H_0$ the set of positive integers $\{10^H \lfloor x_1 \rfloor_H, \dots, 10^H \lfloor x_n \rfloor_H\}$ is a B_h -set.

Proof. Let $\mathbf{c}^1, \dots, \mathbf{c}^m$ be the elements of $\chi_{h,n}$, and apply Proposition 2 and Lemma 4. $\square$

Remark 4. The underlying principle (passing to an infinite subsequence) used in this standard proof is sometimes called Weak König’s Lemma (or WKL_0) by people who work in the area of Logic called *Reverse Mathematics*. Essentially the same argument can be used to give a standard proof of the Sierpinski Diophantine result Theorem 1.

These proofs (both the nonstandard and standard variants) are easier than the original standard proof of Theorem 10. But the approach is not as powerful, in that it does not give you any way of knowing what H_0 to take.

Once one understands this proof, one starts seeing B_h -set constructions everywhere. Here are two examples.

Theorem 11. *Let $h \in \mathbb{N}_+$, and let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}_+$ which is linearly independent over $\mathbb{Q}$. There is a sequence $\{H_n\}_{n \in \mathbb{N}}$ of natural numbers such that for any $n \in \mathbb{N}$ and all $H \geq H_n$, the set $\{10^H \lfloor x_1 \rfloor_H, \dots, 10^H \lfloor x_n \rfloor_H\}$ is a B_h -set.*

Theorem 12. *Let $h \in \mathbb{N}_+$, $X \subseteq \mathbb{R}$ be a finite B_h -set, and $x \in \mathbb{R}$ a real number with infinitely many nonzero digits. Then there is a (finite) $H_0 \in \mathbb{N}$ such that for all $H \geq H_0$, $X \cup \{\lfloor x \rfloor_H\}$ is a B_h -set. It follows that if $x \in \mathbb{R}$ is a real number with infinitely many nonzero digits, then there is a sequence $\{H_n\}_{n \in \mathbb{N}}$ in $\mathbb{N}$ such that $\{\lfloor x \rfloor_{10^{H_n}}\}_{n \in \mathbb{N}}$ is an infinite B_h -set.*

The proofs are left as exercises for the reader.

6. Hat Puzzles

In the previous applications, a nonstandard argument produced an idea that could be turned into a standard proof. In this section we will give an example of an infinitary, but standard, result that can be obtained from a standard finitary one by a simple application of a nonstandard arguments. This is in fact a common theme in applied nonstandard analysis; we will discuss some other examples in the final section.

Start with a simple, well-known puzzle about hats. Hat puzzles constitute a broad class of puzzles in which a set of agents are supposed to guess the color of their own hat just from knowing the colors of the hats of some of the agents around them, as well as knowing the guesses of some of those agents. A descendent of these puzzles is a peculiar 2008 theorem of Hardin and Taylor [14] where the Axiom of Choice is used to predict the future. Discussion of both the hat puzzles described in this section, and many other related puzzles, can be found in Galvin [12] and Hardin and Taylor [15, 16].

For this puzzle:

1. Agents $a_1 \dots a_N$ line up in a queue, a_n facing the back of a_{n+1} .
2. Each agent wears either a white (W) or black (B) hat.
3. In order from back to front, each is asked to guess the color of their own hat. The problem is to find a strategy which minimizes the number of wrong answers.⁷

Some conditions on what is meant by a *strategy* are:

1. Each agent a_n can see the hats on agents a_i for $i > n$, but not their own hat.
2. Agent a_n can hear a_i 's response for $i < n$, and also hear whether a_i guessed correctly.
3. All the agents agree on a strategy before queuing.

In other words, the agents all agree on what they will do when it is their turn to guess, and while this might depend on their position in the queue, the hats they have seen and the guesses and answers they have heard, it should *only* depend on this information, and the strategy should be agreed to before the agents have this information.

Unsurprisingly, there is no strategy which assures that a_1 will always get the right answer. Remarkably, there *is* one that ensures that everyone *else* will guess correctly.

Theorem 13. *There is a strategy so that for all $n > 1$, a_n guesses correctly in this finite hat problem.*

As this is an entertaining puzzle, the proof is left to the reader. Note that the first two conditions together mean that when it is agent a_n 's turn to guess, they will know the hat colors of all the agents except themselves.

We can now ask the natural question: what if there are infinitely many agents in the queue?

To answer this question, we begin by formalizing such puzzles.

Let T be a set which represents the set of agents but can also be interpreted as a position or a timeline. For the finite hat problem, $T = \{1, 2, \dots, N\}$. Let S be another set, a set of *outcomes*; in the hat problem, $S = \{W, B\}$. Write ${}^T S$ for the set of all functions from T into S . (So any element of ${}^T S$ represents a possible configuration of hats worn by the queued agents.)

A *strategy* (for T and S) is then a function

$$\lambda: {}^T S \times T \rightarrow S$$

such that for all $f, g \in {}^T S$ and $t \in T$, if $f = g$ on $T - \{t\}$ then $\lambda(f, t) = \lambda(g, t)$. (In other words, $\lambda(f, t)$ does not depend on the value of f at t .) Denote by $\mathbb{S}(T, S)$ the set of all such strategies. Agent $t \in T$ employs the strategy by observing $f \in {}^T S$

⁷This puzzle is often framed as prisoners being led out into the prison yard and being told that if their guess is wrong they will be executed. The author is not aware of any prison system where this is normal practice.

(giving any value at all to $f(t)$) and then guessing $\lambda(f, t)$. It is easy to see that this agrees with the informal notion of strategy above.

Theorem 13 can now be expressed as a first-order sentence:

$$(\forall N \in \mathbb{N})(\exists \lambda_N \in \mathbb{S}(\{1, \dots, N\}, \{W, B\})) \\ ((\forall f \in {}^{1, \dots, N}\{W, B\})(\forall t \neq 1)[\lambda(f, t) = f(t)]). \quad (4)$$

Now, suppose we have countably many agents $a_n, n = 1, 2, \dots$ lined up as before. We can now prove the following.

Theorem 14. *There is a strategy so that for all $n > 1$, a_n guesses correctly in this infinite hat problem.*

Proof. Let $T = \{a_n : n \in \mathbb{N}\}$ be the set of agents, indexed by $\mathbb{N}$. Moving to the nonstandard universe, T extends to a set *T , indexed by ${}^*\mathbb{N}$, where $a_n \in {}^*T$ for all standard $n \in \mathbb{N}$. Let $N \in {}^*\mathbb{N}$ be an arbitrary infinite integer.

By transfer of Equation (4),

$$(\exists \lambda_N \in {}^*\mathbb{S}(\{1, \dots, N\}, \{W, B\}))((\forall f \in {}^{1, \dots, N}\{W, B\})(\forall t \neq 1)[\lambda(f, t) = f(t)]).$$

In other words, there is a (nonstandard) strategy for the nonstandard puzzle with agents numbered from 1 to N . The strategy for the standard puzzle (with agents T) is then:

1. Start with the nonstandard strategy that exists for the hyperfinite set $\{a_1, \dots, a_N\}$.
2. Standard agents a_n , with $n \in \mathbb{N}$ should employ this strategy to guess W or B when their turn comes.
3. Ignore all the infinite agents, since they do not exist anyway.

Since this strategy creates a correct guess for all n with $1 < n \leq N$, and since ${}^*f(n) = f(n)$ for any $f \in {}^T\{W, B\}$ and $n \in \mathbb{N}$, it will be correct for $1 < n \in \mathbb{N}$. This proves Theorem 14. $\square$

Remark 5. This solution to the infinitary hat puzzle is not exactly new; the puzzle has been solved using a strategy based on a nonprincipal ultrafilter. This is roughly the same solution as given here, as it is well-known that every ultrafilter corresponds to an element of ${}^*\mathbb{N}$, and vice versa. However, the nonstandard approach has the advantage of being straightforward, and requiring little by way of verification.

Nonprincipal ultrafilters are nonconstructive objects, and so this ‘strategy’ will not be satisfying for someone who wants a concrete algorithm. It does not seem possible to remove the ultrafilter from the solution. The relationship of some hat puzzles to some kinds of ultrafilters is discussed in [16].

7. Final Thoughts

We want to conclude by giving at least some mention to two general principles.

7.1. The “Automatically Uniform” Property of Nonstandard Analysis

That is:

If *E has property P then E has property P uniformly.

This is not a real theorem as stated, but more of a common phenomenon (though there are ways to make the statement precise). An example that might be of interest to the reader is a remarkable theorem of Keller [24], given as Theorem 15 below.

Theorem 15. *Let $\mathcal{V}$ be a variety of groups, and suppose every $G \in \mathcal{V}$ is amenable. Then $\mathcal{V}$ is uniformly amenable.*

Recall that a group G is *amenable* if there is an invariant finitely-additive probability measure on G . (This is one of a nimety of definitions!) Since group operations are all first-order, if G is a group then so is *G . If *G is amenable then so is G , since subgroups of amenable groups are amenable. Keller’s observation was that the converse need not be the case, that the nonstandard extension of an amenable group need not be amenable. He therefore called a group *uniformly amenable* if *G is amenable, found an equivalent standard characterization of uniform amenability (in terms of a thing called the Følner condition), and extended the definition to varieties of groups using the latter.

In this way an interesting new standard concept was first discovered in a natural way using a nonstandard principle.

7.2. The Ability to Treat Infinite Sets as Finite

Michael Richter once described nonstandard analysis as

the art of making infinite sets finite by extending them.

The Transfer Principle lets us apply finite combinatorics to hyperfinite objects, and this has been used many times to reveal connections between discrete and continuous results, giving some affirmative answers to the title question in Lázló Lóvász’s paper “Discrete and Continuous: Two sides of the same?”

For example, one can show that the Riesz Representation Theorem from measure theory is ‘really’ the Farkas Lemma for a system of linear inequalities [45, 47]; that the existence of measurable selections of random close sets can be extracted from a finite Marriage Lemma argument [46]; and that the invariance of some measures under local homomorphisms or isometries can be verified by looking at permutation groups [44].

Finally, it can happen that the obvious hyperfinite analogue of a finite result provides an interesting infinitary result with no clear standard counterpart. One example is Peter Loeb’s remarkable ‘discrete Lyapunov Theorem’ [28], which applies Steinitz’s 1913 result about reordering vectors in $\mathbb{R}^n$ to minimize partial sums, to give a result about nearly atomless measures. There ought to be a way to turn this into a standard proof of the Lyapunov Theorem, but this has not been done yet, and mathematical economists interested in this area of convexity theory have opted instead to just remain in the nonstandard universe [26].

Every result in discrete mathematics has, through Transfer, an infinitary counterpart, and almost none have been mined as thoroughly as we believe is possible.

Acknowledgement. The author is grateful to the organizers of Integers 2025 for inviting me to this delightful conference, and to my colleague Mel Nathanson for introducing me to so many interesting mathematical problems of which I was previously unaware.

References

- [1] A. Abian, On solvability of infinite systems of polynomial equations over a finite ring, *Bull. Soc. Math. Belg.* **46** (1972/73), 33–37.
- [2] A. Abian and E. Eslami, Solvability of infinite systems of infinite linear equations, *Bol. Soc. Mat. Mexicana* **27** (2) (1982), 57–64.
- [3] L. O. Arkeryd, N. J. Cutland, and C. W. Henson (eds.), *Nonstandard Analysis*, volume 493 of NATO Advanced Science Institutes Series C, Kluwer, Dordrecht, 1997.
- [4] S. Banach, Sur les fonctionelles linéaires, *Studia Math.* **1** (1929), 223–239.
- [5] S. Banach, *Théorie des Opérations Linéaires*, Chelsea, New York, 1932. Translated and reprinted as *Theory of Linear Operators*, Dover Publications, Mineola, NY, 1987.
- [6] F. Cucker and A. G. Corbalan, An alternate proof of the continuity of the roots of a polynomial, *Amer. Math. Monthly* **96** (1989), 342–345.
- [7] D. R. Curtiss, On Kellogg’s Diophantine problem, *Amer. Math. Monthly* **29** (1922), 380–387.
- [8] N. J. Cutland (ed.), *Nonstandard Analysis and its Applications*, volume 10 of London Mathematical Society Student Texts, Cambridge University Press, Cambridge, 1988.
- [9] M. Davis, *Applied Nonstandard Analysis*, 2nd edition, Dover, Mineola, 2005.
- [10] M. Di Nasso, I. Goldbring, and M. Lupini, *Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory*, Lecture Notes in Math. 2239, Springer, Cham, 2019.
- [11] V. Fowler, ISU professor asks, ‘Why not blow up the moon?’, *Des Moines Register*, March 3, 1991, p. 11.
- [12] F. Galvin, Problem 5348, *Amer. Math. Monthly* **72** (1965), 1019.

- [13] H. Hahn, Über linearer Gleichungssysteme in linearer Raumen, *J. Reine Angew. Math.* **157** (1927), 214–229.
- [14] C. S. Hardin and A. D. Taylor, A peculiar connection between the axiom of choice and predicting the future, *Amer. Math. Monthly* **115** (2) (2008), 91–96.
- [15] C. S. Hardin and A. D. Taylor, An introduction to infinite hat problems, *Math. Intelligencer* **30** (4) (2008), 20–25.
- [16] C. S. Hardin and A. D. Taylor, *The Mathematics of Coordinated Inference: A Study of Generalized Hat Problems*, Springer, Cham, 2013.
- [17] G. H. Hardy, *Orders of Infinity. The Infinitärkalkül of Paul du Bois-Reymond*, Cambridge University Press, Cambridge, 1910.
- [18] G. Harris and C. Martin, The roots of a polynomial vary continuously as a function of the coefficients, *Proc. Amer. Math. Soc.* **100** (1987), 390–392.
- [19] E. Helly, Über linearer Funktionaloperationen, *Sitzungsber. der math. Naturwiss. Klasse der Akad. der Wiss. (Wien)* **121** (1912), 265–297.
- [20] E. Helly, Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten, *Monatsh. für Math. Phys.* **31** (1921), 60–91.
- [21] M. Henriksen and J. R. Isbell, On the continuity of the real roots of an algebraic equation, *Proc. Amer. Math. Soc.* **4** (1953), 431–434.
- [22] K. Hirose, Continuity of the roots of a polynomial, *Amer. Math. Monthly* **127** (2020), 359–363.
- [23] R. Jin, Characterizing the structure of $A + B$ when $A + B$ has small upper Banach density, *J. Number Theory* **130** (8) (2010), 1785–1800.
- [24] G. Keller, Amenable groups and varieties of groups, *Illinois J. Math.* **16** (1972), 257–268.
- [25] O. D. Kellogg, On a Diophantine problem, *Amer. Math. Monthly* **28** (1921), 300–303.
- [26] M. A. Khan and K. P. Rath, The Shapley-Folkman theorem and the range of a bounded measure: an elementary and unified treatment, *Positivity* **17** (3) (2013), 381–394.
- [27] T. Lindström, An invitation to nonstandard analysis, in *Nonstandard Analysis and its Applications*, Cambridge University Press, Cambridge, 1988, 1–105.
- [28] P. A. Loeb, A combinatorial analog of Lyapunov’s theorem for infinitesimally generated atomic vector measures, *Proc. Amer. Math. Soc.* **39** (1973), 585–586.
- [29] W. A. J. Luxemburg, Two applications of the method of construction by ultrapowers to analysis, *Bull. Amer. Math. Soc.* **68** (1962), 416–419.
- [30] W. A. J. Luxemburg, Reduced powers of the real number system and equivalents of the Hahn-Banach extension theorem, in *Int. Sympos. on the Application of Model Theory to Alg. Anal. and Probability*, Holt, Rinehart and Winston, New York, 1969, 123–137.
- [31] M. Marden, *Geometry of Polynomials*, Math. Surveys Monogr. **3**, Amer. Math. Soc., Providence, 1966.
- [32] M. McKinzie and C. Tuckey, Higher trigonometry, hyperreal numbers, and Euler’s analysis of infinities, *Math. Mag.* **74** (5) (2001), 339–368.

- [33] L. Narici and E. Beckenstein, The Hahn-Banach Theorem: the life and times, *Topology Appl.* **77** (1997), 193–211.
- [34] M. B. Nathanson, Weighted real Egyptian numbers, *Funct. Approx. Comment. Math.* **60** (2019), 15–166.
- [35] M. B. Nathanson and D. A. Ross, Continuity of the roots of a polynomial, *Comm. Algebra* **52** (6) (2024), 2509–2518.
- [36] M. B. Nathanson and D. A. Ross, Linear equations and multiplicative polynomial equations in infinitely many variables, *Comm. Algebra* **53** (6) (2024), 2302–2318.
- [37] M. B. Nathanson, Finitely many implies infinitely many, *J. Number Theory* **278** (2026), 115–124.
- [38] M. B. Nathanson, $\mathbf{Q}$ -independence and the construction of B_h -sets of integers and lattice points, *Discrete Math.* **349** (2) (2026), Paper No. 114726.
- [39] A. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces*, Academic Press, 1973.
- [40] D. Pincus, The strength of the Hahn-Banach Theorem, in *Victoria Symposium on Nonstandard Analysis*, Lecture Notes in Mathematics 369, Springer-Verlag, Berlin, 1974, 203–248.
- [41] F. Riesz, *Les Systèmes d'Équations Linéaires à une Infinité d'Inconnues*, Gauthier-Villars, Paris, 1913.
- [42] A. Robinson, *Non-standard Analysis*, 2nd edition, North-Holland, Amsterdam, 1974.
- [43] Q. I. Rahman and G. Schmeisser, *Analytic Theory of Polynomials*, Clarendon Press, Oxford, 2002.
- [44] D. A. Ross, Measures invariant under local homeomorphisms, *Proc. Amer. Math. Soc.* **102** (4) (1988), 901–905.
- [45] D. A. Ross, Yet another short proof of the Riesz representation theorem, *Math. Proc. Cambridge Philos. Soc.* **105** (1) (1989), 139–140.
- [46] D. A. Ross, Selectionable distributions for a random set, *Math. Proc. Cambridge Philos. Soc.* **108** (2) (1990), 405–408.
- [47] D. A. Ross, Pushing down infinite Loeb measures, *Math. Scand.* **104** (1) (2009), 108–116.
- [48] D. A. Ross, Egyptian fractions, nonstandard extensions of $\mathbb{R}$, and some Diophantine equations without many solutions, in *Combinatorial and Additive Number Theory III*, Springer, Cham, 2020.
- [49] D. A. Ross, Egyptian fractions on groups, in *Combinatorial and Additive Number Theory IV*, Springer, Cham, to appear.
- [50] W. Sierpinski, Sur les décompositions de nombres rationnels en fractions primaires, *Mathesis* **65** (1956), 16–32.
- [51] T. Skolem, Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen, *Fund. Math.* **23** (1) (1934), 150–161.
- [52] J. Valente, Hate winter? Here's a scientist's answer: Blow up the moon, *The Wall Street Journal*, April 22, 1991.
- [53] H. Whitney, *Complex Analytic Manifolds*, Addison-Wesley, Reading, 1972.