An Invitation to Nonstandard Analysis and its Recent Applications

Isaac Goldbring and Sean Walsh

Introduction

Nonstandard analysis has been used recently in major results, such as Jin's sumset theorem in additive combinatorics and Breuillard–Green–Tao's work on the structure of approximate groups. However, its roots go back to Robinson's formalization of the infinitesimal approach to calculus. After first illustrating its very basic uses in calculus in "Calculus with Infinitesimals," we go on to highlight a selection of its more serious achievements in "Selected Classical and Recent Applications," including the aforementioned work of Jin and Breuillard–Green–Tao. After presenting a simple axiomatic approach to nonstandard analysis in "Axioms for Nonstandard Extensions" we examine Jin's theorem in more detail in "The Axioms in Action: Jin's Theorem." Finally, in "The Ultraproduct Construction" we discuss how these axioms can be justified

Isaac Goldbring is an associate professor in the Department of Mathematics at UC Irvine. His email address is `isaac@math.uci.edu`.

Sean Walsh is an associate professor in the Department of Philosophy at UCLA. His email address is `walsh@ucla.edu`.

DOI: https://doi.org/10.1090/noti1895

with a particular concrete construction (akin to the verification of the axioms for the real field using Dedekind cuts or Cauchy sequences), and in "Other Approaches" we compare our axiomatic approach to other approaches. While brief, our hope is that this survey can quickly give the reader a sense of both how nonstandard methods are being used today and how these methods can be rigorously presented and justified.

Calculus with Infinitesimals

Every mathematician is familiar with the fact that the founders of calculus such as Newton and Leibniz made free use of *infinitesimal* and *infinite* quantities, e.g., in expressing the derivative $f'(a)$ of a differentiable function f at a point a as being the real number that is infinitely close to $\frac{f(a+\delta)-f(a)}{\delta}$ for every nonzero infinitesimal "quantity" δ. Here, when we say that δ is infinitesimal, we mean that $|\delta|$ is less than r for every positive real number r.

Of course, the mathematical status of such quantities was viewed as suspect and the entirety of calculus was put on firm foundations in the nineteenth century by the likes of Cauchy and Weierstrass, to name a few of the more significant figures in this well-studied part of the history of mathematics. The innovations of their "ϵ-δ method" (much to the chagrin of many real analysis students today) allowed one to give rigor to the naïve arguments of their predecessors.

In the 1960s, Abraham Robinson realized that the ideas and tools present in the area of logic known as *model theory* could be used to give precise mathematical meanings to infinitesimal quantities. Indeed, one of Robinson's stated aims was to rescue the vision of Newton and Leibniz ([20]). While there are complicated historical and philosophical questions about whether Robinson succeeded entirely in this (cf. [5]), our goal is to discuss some recent applications of the methods Robinson invented.

To this end, let us turn now to the basics of Robinson's approach. On this approach, an infinitesimal δ is merely an element of an ordered field $\mathbf{R}^*$ properly containing the ordered field $\mathbf{R}$ of real numbers. While such *nonarchimedean* fields were already present in, for example, algebraic number theory, the new idea that allowed one to correctly apply the heuristics from the early days of calculus was that $\mathbf{R}^*$ is logically similar to $\mathbf{R}$, in that any elementary statement true in the one is also true in the other. (Formally, elementary statements are defined using first-order logic: see "Axioms for Nonstandard Extensions" for a self-contained presentation.) A particular feature of this approach is that one functorially associates to every function $f : \mathbf{R} \to \mathbf{R}$ an extension $f : \mathbf{R}^* \to \mathbf{R}^*$ (and similarly for relations). In this light, $\mathbf{R}$ is referred to as the *standard field of real numbers* whilst $\mathbf{R}^*$ is referred to as *a nonstandard field of real numbers* or a *hyperreal field*.

To illustrate this method, let us say that two elements a and b of $\mathbf{R}^*$ are *infinitely close* to one another, denoted $a \approx b$, if their difference is an infinitesimal. Then one has:

Theorem 1. *Suppose that $f : \mathbf{R} \to \mathbf{R}$ is a function and $a \in \mathbf{R}$. Then f is continuous at a if and only if: whenever $b \in \mathbf{R}^*$ and $a \approx b$, then $f(a) \approx f(b)$. Likewise, f is differentiable at a with derivative L if and only if: whenever b is a nonzero infinitesimal in $\mathbf{R}^*$, one has that $\frac{f(a+b)-f(a)}{b} \approx L$.*

Proof. We only discuss the continuity statement, since the differentiability statement is entirely analogous. Suppose that the assumption of the "if" direction holds and fix $\epsilon > 0$ in $\mathbf{R}$. In $\mathbf{R}^*$, by choosing an infinitesimal $\delta > 0$, the following elementary statement about ϵ is true: "there is $\delta > 0$ such that, for all b, if $|a-b| < \delta$, then $|f(a) - f(b)| < \epsilon$." The logical similarity mentioned above then implies that this elementary statement about ϵ is also true in $\mathbf{R}$. But this is precisely what is needed to prove that f is continuous at a. The other direction is similar. □

Every finite element a of $\mathbb{R}^*$ is infinitely close to a unique real number, called its *standard part*, denoted $\mathrm{st}(a)$. Thus, the first part of the above theorem reads: f is continuous at a if and only if: whenever $\mathrm{st}(b) = a$, then $\mathrm{st}(f(b)) = f(a)$. In these circumstances, one can then write $b = a + \delta x$ and $f(b) = f(a) + \delta y$ where $\delta x, \delta y$ are infinitesimals. In his elementary calculus textbook based on nonstandard methods, Keisler pictured this as a "microscope" by which one can zoom in and study the local behavior of a function at a point.

Besides extensionally characterizing familiar concepts such as continuity and differentiability, the use of infinitesimals can help to abbreviate proofs. To illustrate, let us prove the following, noting how we can avoid the usual "$\frac{\epsilon}{2}$" argument and instead appeal to the simple fact that finite sums of infinitesimals are again infinitesimal:

Corollary 2. *Suppose that $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ are both continuous at $a \in \mathbf{R}$. Then $f + g$ is also continuous at a.*

Proof. Consider $b \approx a$. By assumption, we have $f(b) \approx f(a)$ and $g(b) \approx g(a)$. Then since the sums of two infinitesimals is infinitesimal, we have

$$(f+g)(b) = f(b) + g(b) \approx f(a) + g(a) = (f+g)(a).$$

Since b was arbitrary, we see that $f + g$ is continuous at a. □

Selected Classical and Recent Applications

The primary reason for the contemporary interest in nonstandard analysis lies in its capacity for proving new results. In this section, we briefly describe a handful of the more striking results that were first proven using nonstandard methods. We remark that nonstandard methods have proven useful in nearly every area of mathematics, including algebra, measure theory, functional analysis, stochastic analysis, and mathematical economics to name a few. Here, we content ourselves with a small subset of these application areas.

Bernstein-Robinson theorem on invariant subspaces. The famous *Invariant Subspace Problem* asks whether every bounded operator T on a separable Hilbert space H has a T-invariant closed subspace besides $\{0\}$ and H. In the 1930s, von Neumann proved that *compact* operators have nontrivial closed invariant subspaces. (Recall an operator is compact if it is the norm-limit of finite-rank operators.) Little progress was made on the Invariant Subspace Problem until:

Theorem 3 (Bernstein-Robinson [3]). *If T is a* polynomially compact *operator on H (meaning that there is a nonzero polynomial $p(Z) \in \mathbb{C}[Z]$ such that $p(T)$ is a compact operator), then T has a nontrivial closed invariant subspace.*

One of the main ideas of the proof is to use the nonstandard version of the basic fact that operators on finite-dimensional spaces have upper-triangular representations to find many hyperfinite-dimensional subspaces of H^* which are T^*-invariant. (Upon seeing the Berenstein-Robinson theorem, Halmos proceeded to give a proof that did not pass through nonstandard methods and the two papers were published together in the same volume.) We should point out that the Berenstein-Robinson theorem was later subsumed by Lomonosov's theorem from 1973, which says that an operator that commutes with a nonzero compact operator satisfies the conclusion of the Invariant Subspace Problem.

Asymptotic cones. In [11], Gromov proved the following theorem, which is one of the deepest and most beautiful theorems in geometric group theory:

Theorem 4. *If G is a finitely generated group of polynomial growth, then G is virtually solvable.*

Here, a finitely generated group has *polynomial growth* if the set of group elements that can be written as a product of at most d generators and their inverses grows polynomially in d. The polynomial growth condition is a geometric condition describing the growth of the sizes of balls centered at the identity in the Cayley graph of the group and the amazing fact represented in this theorem is that this geometric condition has serious algebraic consequences. (The converse of the theorem is also true and much easier to prove.)

A key construction in the proof of Gromov's theorem is that of an *asymptotic cone* of a metric space. Roughly speaking, an asymptotic cone of a metric space is the result of looking at the metric space from "very far away" retaining only the large-scale geometry of the space. The prime example of this phenomenon is that an asymptotic cone of the discrete space $\mathbf{Z}$ is the continuum $\mathbf{R}$.

In [8], van den Dries and Wilkie used the nonstandard perspective to give a much cleaner account of the asymptotic cone construction. Indeed, given a metric space (X, d), a fixed point $x_0 \in X$, and an infinite hyperreal number $R \in \mathbf{R}^*$, one can look at the subspace

$$X_R := \left\{ x \in X^* \ : \ \frac{d(x, x_0)}{R} \text{ is finite} \right\}.$$

One can then place the metric d_R on X_R given by $d_R(x, y) := \mathrm{st}(\frac{d(x,y)}{R})$. When X is the Cayley graph associated to a group, then the polynomial growth condition on Γ allows one to find R so that X_R is locally compact; the verification of this and other important properties of X_R are very clear from the nonstandard perspective. While Gromov's original proof did not use nonstandard methods, the nonstandard perspective on asymptotic cones has now in fact become the one that is presented in courses and textbooks on the subject. See, for instance, [18].

Jin's Sumset Theorem. In additive combinatorics, the focus is often on densities and structural properties of subsets of $\mathbf{N}$. Given $A \subseteq \mathbf{N}$, we define the *Banach density of A* to be $\mathrm{BD}(A) := \lim_{n\to\infty} \max_{x \in \mathbf{N}} \frac{|A \cap [x, x+n)|}{n}$. If $\mathrm{BD}(A) > 0$, then we think of A as a "large" subset of $\mathbf{N}$. An important structural property of sets of natural numbers is that of being *piecewise syndetic*, where A is piecewise syndetic if there is $m \in \mathbf{N}$ such that $A + [0, m]$ contains arbitrarily long intervals. Renling Jin [14] used nonstandard analysis to prove the following:

Theorem 5. *If $A, B \subseteq \mathbf{N}$ both have positive Banach density, then $A + B$ is piecewise syndetic.*

Jin's theorem has paved the way for further applications of nonstandard methods in additive combinatorics, which is now a very active area; see the monograph [6]. In "The Axioms in Action: Jin's Theorem," we discuss the proof of Jin's theorem, using the formal framework for nonstandard analysis that we set out in "Axioms for Nonstandard Extensions."

The structure of approximate groups. Fix $K \geq 1$. A symmetric subset A of a finite group G is said to be a *K-approximate subgroup of G* if $A \cdot A$ is contained in K (left) translates of A. Approximate subgroups are generalizations of subgroups (since a 1-approximate group is simply a subgroup of G). The *Freiman theorem for abelian groups* classifies K-approximate subgroups of abelian groups (they are, in some sense, built from generalized arithmetic progressions and group extensions). It was an open problem whether or not a similar result held for K-approximate subgroups of arbitrary (not necessarily abelian) groups.

In [13], using sophisticated methods from model theory, Hrushovski made a major breakthrough in the above so-called nonabelian Freiman problem. While he did not completely solve the problem, his breakthrough paved the way for the work of Breuillard, Green, and Tao [4], who reduced the use of sophisticated model theory in favor of mere nonstandard analysis (while using more intricate combinatorics) and, in the process, succeeded in providing a complete classification of arbitrary approximate groups.

The key insight of Hrushovski (which is also present in the Breuillard–Green–Tao work) is that a nonstandard (infinite) approximate group can be "modelled" by a locally compact group, which, by the structure theory of locally compact groups, can in turn be modeled by finite-dimensional Lie groups. (This is not unrelated to the use of nonstandard methods in the proof of Gromov's theorem described above.) Thus, in attacking this problem of finite combinatorics, one can use the infinitary tools from differential geometry and Lie theory. Blurring the distinction between the finite and the infinite is a cornerstone of many applications of nonstandard methods, allowing one to leverage tools from either side of the divide whenever it proves convenient. Indeed, this is present even in the characterization of differentiability from Theorem 1, where the infinitary process of taking a limit is replaced by the discretized finite quotient.

We mention in passing that the structure theory of locally compact groups, due to Gleason, Montgomery, and Zippin, solved the fifth problem of Hilbert. Hirschfeld [12] used nonstandard analysis to give a conceptually simpler account of this solution. The first-named author generalized Hirschfeld's account to solve the local version of Hilbert's fifth problem [10]; this was, in turn, a part of the Breuillard–Green–Tao work mentioned above.

Axioms for Nonstandard Extensions

In this section, we describe an approach to nonstandard analysis using three axioms (NA1)-(NA3). This follows and elaborates on the treatment in the appendix of the first-named author's earlier coauthored work [7]. However, when specialized to the real numbers, axioms (NA1)-(NA2) are very similar to Goldblatt's [9] presentation of nonstandard methods.

The approach through (NA1)-(NA3) should remind the reader of the beginning of measure theory, where one sets out carefully a class of sets that form the building blocks of the subject-matter one is interested in. The most basic sets are those built into what we call the *structures* in (NA1)—and then we expand upon this class with the *definable sets* of axiom (NA2). This is similar to how, in measure theory, one might pass from basic half-opens on the real line to the σ-algebra generated by them. However, in nonstandard analysis, we build up more complicated sets from basic sets simultaneously in a structure and a nonstandard extension, similar to how one might look at varieties both over a field and over various of its field extensions. Finally, for our applications, we need a final axiom (NA3) that ensures that our class of definable sets in the nonstandard extension has certain compactness properties.

Structures and axiom (NA1).

Definition 6. A *structure* $\mathbf{S} = ((S_i), (R_j))$ is a collection of sets $(S_i \;:\; i \in I)$, often called the *basic sets* of $\mathbf{S}$, together with a collection $(R_j \;:\; j \in J)$ of *distinguished relations* on the basic sets, that is, for each $j \in J$, there are $i(1), \ldots, i(n) \in I$ such that $R_j \subseteq S_{i(1)} \times \cdots \times S_{i(n)}$. Distinguished relations are also called *primitives* of $\mathbf{S}$.

Later, it will be convenient to speak of the *complete structure on* (S_i), which is simply the structure with basic sets (S_i) and where we take *all* relations as the basic relations.

We can now state the first axiom of nonstandard extensions, which simply explains what makes them extensions:

(NA1) Each basic set S_i is extended to a set $S_i^* \supseteq S_i$ and, to each distinguished relation R_j as above, we associate a corresponding relation

$$R_j^* \subseteq S_{i(1)}^* \times \cdots \times S_{i(n)}^*$$

whose intersection with $S_{i(1)} \times \cdots \times S_{i(n)}$ is the original relation R_j.

Definable sets and axiom (NA2). Let $\mathbf{S}$ be a structure. Some notation: given $\bar{i} := (i(1), \ldots, i(n))$, we set $S_{\bar{i}} := S_{i(1)} \times \cdots \times S_{i(n)}$. It will also be necessary to declare $S_\varnothing$ to be a one-element set. For the sake of readability, if $\bar{i}$ and $\bar{j}$ are two finite sequences, then we write $\bar{i}\,\bar{j}$ for the concatenation of the two sequences; if $\bar{i} = (i)$, then we simply write $i\,\bar{j}$ and similarly for when $\bar{j}$ is a one-element sequence.

We define the collection of $\mathbf{S}$-*definable sets* to be the Boolean algebras $\mathcal{D}_{\mathbf{S}}(\bar{i})$ (or simply $\mathcal{D}(\bar{i})$ if $\mathbf{S}$ is clear from context) of subsets of $S_{\bar{i}}$ with the following properties:

(1) $\varnothing, S_{\bar{i}} \in \mathcal{D}(\bar{i})$.
(2) For any i, $\{(x, y) \in S_i \times S_i \;:\; x = y\} \in \mathcal{D}(i, i)$.
(3) If $a \in S_i$, then $\{a\} \in \mathcal{D}(i)$.
(4) If $R_j \subseteq S_{\bar{i}}$ is a basic relation, then $R_j \in \mathcal{D}(\bar{i})$.
(5) If $A \in \mathcal{D}(\bar{i})$, then $A \times S_j \in \mathcal{D}(\bar{i}\,j)$ and $S_j \times A \in \mathcal{D}(j\,\bar{i})$.
(6) If $A \in \mathcal{D}(\bar{i})$ and $\vec{i} = \overline{i_1}\; j\; \overline{i_2}$ and $\pi \;:\; S_{\bar{i}} \to S_{\overline{i_1}\,\overline{i_2}}$ is the canonical projection, then $\pi(A) \in \mathcal{D}(\overline{i_1}\,\overline{i_2})$.

While the closure properties of a Boolean algebra encode propositional operations such as conjunction and disjunction, the additional postulates on the definable sets encode the operations of predicate logic: the identity relation is homophonically encoded in (2), while the last condition (6) encodes existential quantification since $\overline{a_1}\ \overline{a_2}$ is in the projection $\pi(A)$ iff there is b in S_j with $\overline{a_1}\ b\ \overline{a_2}$ belonging to A.

Now, if $D \in \mathcal{D}(\bar{i})$ is an $\mathbf{S}$-definable set, then define the $\mathbf{S}^*$-definable set D^* simply by replacing all instances of basic relations R_j used in the construction of D by R_j^*. Note that not every $\mathbf{S}^*$-definable set is of the form D^* for an $\mathbf{S}$-definable set, e.g., $\{a\}$ for $a \in S_i^* \backslash S_i$. And we will soon see that such elements exist.

We can now explain the second axiom of nonstandard analysis, the so-called *transfer principle*, which in the terminology of definable sets simply says that the property of basic relations from (NA1) also holds for arbitrary $\mathbf{S}$-definable sets:

(NA2) For any $\mathbf{S}$-definable set $D \in \mathcal{D}(\bar{i})$, we have $D^* \cap S_{\bar{i}} = D$.

Axiom (NA2) states precisely what it means for the nonstandard extension to be logically similar to the original structure. It is saying that any statement about elements of the original structure that can be phrased in terms of definable sets also holds within the nonstandard extension.

Compactness and richness (NA3). Our final axiom asks that our nonstandard extensions contain sufficiently many "ideal" elements analogous to the infinitesimal elements added to $\mathbf{R}^*$ in "Calculus with Infinitesimals." Here is the precise statement:

The structure $\mathbf{S}$ is said to be *countably rich* if for all $\vec{i}$ and every countable family $(X_m)_{m\in\mathbb{N}}$ of elements of $\mathcal{D}(\bar{i})$ with the finite intersection property, we have $\bigcap_m X_m \neq \emptyset$. Here the "finite intersection property" means that $X_{m_1} \cap \cdots \cap X_{m_k} \neq \emptyset$ for all $m_1, \ldots, m_k \in \mathbb{N}$.

Countable richness can be seen as a logical compactness property when stated in its contraposed form: if $\mathbf{S}$ is countably rich and an $\mathbf{S}$-definable set $X \in \mathcal{D}(\bar{i})$ is covered by countably many $\mathbf{S}$-definable sets $X_m \in \mathcal{D}(\bar{i})$, then X is already covered by finitely many of these X_m.

Here is the final axiom of nonstandard extensions:

(NA3) $\mathbf{S}^*$ is countably rich.

Note that, as an ordered set, $\mathbf{R}$ is not countably rich, since $\bigcap_n (n, +\infty) = \emptyset$. Thus our initial structure $\mathbf{S}$ will usually not be rich, which is why we consider now an extension $\mathbf{S}^*$ of $\mathbf{S}$ such that (NA1), (NA2), and (NA3) hold, so $\mathbf{S}^*$ is countably rich. One consequence is that if X is countable and infinite, then X^* will have elements that are not in X.

Richness is the precise formulation of asking that our nonstandard extension have certain ideal elements analogous to the infinitesimal elements of $\mathbf{R}^*$. For example, if G is a topological group, we will want there to be elements $a \in G^*$ that are infinitely close to the identity element of G in the sense that $a \in U^*$ for every neighborhood U of the identity in G. If G is first countable, then countable richness will ensure the existence of such elements.

We should note that, for certain applications of nonstandard methods, countable richness is not a strong enough assumption. For example, we may want the analogous version of countable richness to hold for families of definable sets indexed by the real numbers rather than the natural numbers. (Such richness is desirable, for example, in applications to combinatorial number theory; see [6].) In this latter scenario, one would then assume that the nonstandard extension is $\mathfrak{c}$-rich, where $\mathfrak{c}$ is the cardinality of the real numbers. One can of course speak of greater assumptions of richness, but we will refrain from dwelling too much on this point.

This concludes the list of our axioms for nonstandard analysis. We remark that we could have also added distinguished functions in our definition of a structure, and then a requirement in (NA2) for extensions of distinguished functions. But working with graphs of functions, it is routine to see that such a set-up can be encoded in a purely relational structure as above.

We now stress: *all applications of nonstandard analysis can proceed from the above three axioms alone.*

Power sets and internal sets. The reader already familiar with nonstandard analysis will notice that we have yet to speak about a notion that is of central importance, namely the notion of internal sets. We remedy that now.

For certain basic sets S of our structure $\mathbf{S}$, we often include also its power set $\mathcal{P}(S)$ as a basic set, and the membership relation $\in_S := \{(x, Y) \in S \times \mathcal{P}(S) : x \in Y\}$ as a primitive. This allows us to quantify over elements of $\mathcal{P}(S)$, and thus gives enormous expressive power. For example, with $\mathbf{R}$ and $\mathcal{P}(\mathbf{R})$ as basic sets, together with the membership relation between them and the ordering on $\mathbf{R}$ as primitives, we can express by an elementary statement the fact that every nonempty subset of $\mathbf{R}$ with an upper bound in $\mathbf{R}$ has a least upper bound in $\mathbf{R}$ (i.e., the completeness of the real line).

Given $Y \in \mathcal{P}(S)$, the formula $v \in_S Y$ (with v a variable ranging over S) defines the subset Y of S, so Y is not only an element in our structure $\mathbf{S}$, but also an $\mathbf{S}$-definable subset of S. In particular, *every* subset of S is now $\mathbf{S}$-definable (while not every subset of $\mathcal{P}(S)$ is $\mathbf{S}$-definable).

Next, let an extension $\mathbf{S}^*$ of $\mathbf{S}$ be given that satisfies (NA1) and (NA2). Then S^* and $\mathcal{P}(S)^*$ are basic sets of $\mathbf{S}^*$, and the star extension $\in_S^*$ of $\in_S$ is among the primitives. There is no reason for the elements of $\mathcal{P}(S)^*$ to be actual subsets of S^*, or for $\in_S^*$ to be an actual membership relation, but we always arrange this to be the case: just replace $P \in \mathcal{P}(S)^*$ by the subset $\{a \in S^* : a \in_S^* P\}$ of S^*. This procedure is traditionally called *Mostowski collapse*, and it identifies $\mathcal{P}(S)^*$ with a subset of $\mathcal{P}(S^*)$.

The subsets of S^* that belong to $\mathcal{P}(S)^*$ via this identification are traditionally called *internal subsets of* S^*. They are in fact exactly the $\mathbf{S}^*$-definable subsets of S^*, as is easily checked.

The Axioms in Action: Jin's Theorem

In this section, we give some details into the proof of Jin's theorem from "Jin's Sumset Theorem."

A few words on $\mathbb{N}^*$. First, it behooves us to give a picture of $\mathbb{N}^*$. We begin by noting that, by transfer, every element N of $\mathbb{N}^*\backslash\mathbb{N}$ is above every element $\mathbb{N}$ in the ordering. Transfer implies that any element N in $\mathbb{N}^*\backslash\mathbb{N}$ has a predecessor in $\mathbb{N}^*$, which we denote by $N-1$; clearly $N-1$ is also infinite, whence we may consider its predecessor $N-2$, another infinite element, and so on. Consequently, N is contained in a copy of the integers (called a $\mathbb{Z}$-chain), the entirety of which is contained in the infinite part of $\mathbb{N}^*$. Note that $2N$ is an infinite number that is contained in a strictly larger $\mathbb{Z}$-chain, and that $\frac{N}{2}$ is an infinite number contained in a strictly smaller $\mathbb{Z}$-chain. Finally, note that if $N<M$ are both infinite, then $\frac{N+M}{2}$ is contained in a $\mathbb{Z}$-chain strictly in-between the $\mathbb{Z}$-chains determined by N and M. We summarize:

Lemma 7. *The set of $\mathbb{Z}$-chains determined by infinite elements of $\mathbb{N}^*$ is a dense linear order without endpoints.*

By an *interval* $I \subseteq \mathbb{N}^*$, we mean a subset of the form

$$(a,b) := \{x \in \mathbb{N}^* \ : \ a < x < b\}$$

for $a<b$ in $\mathbb{N}^*$. The interval is said to be infinite if $b-a$ is infinite.

Given a (finite) interval $I \subseteq \mathbb{N}$, every subset D of I, being itself finite, of course has a cardinality that is an element of $[0,|I|)$. Analogously, given an infinite interval I in $\mathbb{N}^*$ and a definable subset D of I, there is a natural notion of the *definable cardinality* $|D|$ of D, which is an element of $[0,|I|)$. Indeed, we can view the cardinality function $|\cdot|$ as a function on the product of basic sets $\mathcal{P}(\mathbb{N})\times\mathbb{N}\times\mathbb{N}$ given by $|\cdot|(D,a,b) := |D\cap(a,b)|$ if $a<b$ (and some other default value otherwise); the nonstandard extension of this function then assigns a cardinality to definable subsets of intervals in $\mathbb{N}^*$. By transfer, the definable cardinality of an interval is simply its length.

Piecewise syndeticity. Our next order of business is to give a nonstandard description of piecewise syndeticity, which we introduced in "Jin's Sumset Theorem." Given $C \subseteq \mathbb{N}^*$, a *gap* of C on I is a subinterval J of I such that $C\cap J=\emptyset$.

Theorem 8. *$C \subseteq \mathbb{N}$ is piecewise syndetic if and only if there is an infinite interval $I \subseteq \mathbb{N}^*$ such that C^* has only finite gaps on I.*

Proof. We only prove the "if" direction. Suppose that I is an infinite interval such that C^* has only finite gaps on I. We first note that there is $m \in \mathbb{N}$ such that C^* has only gaps of length at most m on I, whence $I \subseteq C^*+[0,m]$. Indeed, if this were not the case, then the set $X_m := \{x \in I \ : \ [x,x+m)\cap C^*=\emptyset\}$ is a nonempty definable subset of $\mathbb{N}^*$ for each m, whence by countable saturation, there is $x\in\bigcap_m X_m$; we then have that $x,x+1,x+2,\ldots \notin C^*$, contradicting that C^* has only finite gaps on I.

Now, given $k\in\mathbb{N}$, let $Y_k := \{x \in \mathbb{N} \ : \ [x,x+k)\subseteq C+[0,m]\}$. By the last paragraph and transfer, we have that $Y_k^* \neq \emptyset$, whence, by transfer again, we have that $Y_k\neq\emptyset$. Since k was arbitrary, this proves that C is piecewise syndetic. □

As an aside, we invite the reader to use the previous theorem to give a quick nonstandard proof that the class of piecewise syndetic subsets of $\mathbb{N}$ is *partition regular*, meaning that if A is piecewise syndetic and $A=B\cup C$, then at least one of B or C is piecewise syndetic. Partition regularity is a crucial notion in Ramsey theory, and the fact that the class of piecewise syndetic sets is partition regular indicates that it is a robust structural notion of largeness.

For $x,y\in\mathbb{N}^*$, we write $x\sim_{\mathbb{N}} y$ if and only if $|x-y|$ is finite. This is an equivalence relation on $\mathbb{N}^*$ and we set $C_{\mathbb{N}}$ for the set of equivalence classes: per Lemma 7, this is just the equivalence class of $\mathbb{N}$ together with the $\mathbb{Z}$-chains. We let $\pi_{\mathbb{N}}:\mathbb{N}^*\to C_{\mathbb{N}}$ denote the canonical projection. Again by Lemma 7, the linear order on $\mathbb{N}^*$ descends to a dense linear order on $C_{\mathbb{N}}$. We can thus restate the previous theorem as: C is piecewise syndetic if and only if $\pi_{\mathbb{N}}(C)$ contains an interval.

Cuts. We now consider a different equivalence relation on $\mathbb{N}^*$. Fix an infinite $N\in\mathbb{N}^*$ and consider instead the relation $\sim_N$ on $\mathbb{N}^*$ given by $x\sim_N y$ if and only if $\frac{|x-y|}{N}$ is infinitesimal. We let C_N denote the set of equivalence classes and $\pi_N:\mathbb{N}^*\to C_N$ the canonical projection. This time, something interesting happens: as a linear order, an initial segment of C_N is isomorphic to the set of positive reals (in particular, the initial segment consists of those x such that $\frac{x}{N}$ is finite, and the isomorphism is given by sending x to $\mathrm{st}(\frac{x}{N})$).

The equivalence relations $\sim_{\mathbb{N}}$ and $\sim_N$ are instances of a more general notion. We call an initial segment U of $\mathbb{N}^*$ a *cut* if U is closed under addition. Note that $\mathbb{N}$ is the smallest cut. Another example of a cut is the cut $U_N := \{K\in\mathbb{N}^* \ : \ \frac{K}{N}$ is infinitesimal$\}$, where N is infinite. As above, for a cut U, defining $x\sim_U y$ if $|x-y|\in U$ yields an equivalence relation on $\mathbb{N}^*$ with set of equivalence classes C_U and projection map π_U. (We chose $\sim_N$ and π_N as opposed to $\sim_{U_N}$ and π_{U_N} for notational cleanliness.) As before, the usual order on $\mathbb{N}^*$ descends to a linear order on C_U.

We now recall a classical theorem of Steinhaus: if E and F are subsets of $\mathbb{R}$ of positive Lebesgue measure, then $E+F$

contains an interval. Given that reals are simply equivalence classes modulo the cut U_N and in proving Jin's theorem we are looking for sums of cuts modulo $U_{\mathbb{N}}$ to contain an interval, it raises the question as to whether or not there is a natural measure on any cut space C_U for which the analogue of Steinhaus' theorem is true and which yields Lebesgue measure in the case of the cut space C_N.

Earlier, we saw that every definable subset of an interval I in $\mathbb{N}^*$ has a definable cardinality. This procedure leads to a natural measure μ_I on the algebra of definable subsets of I given by $\mu_I(D) := \mathrm{st}(\frac{|D|}{|I|})$. The usual Carathéodory extension procedure shows that μ_I can be extended to a σ-additive probability measure on the σ-algebra generated by the definable subsets of I. Indeed, countable richness of the nonstandard extension ensures that the hypotheses of the Carathéodory's extension theorem apply. The resulting measure is called the *Loeb measure* on I with corresponding σ-algebra of *Loeb measurable sets*.

If we are in the situation that $I = [0, N)$ for N infinite and U is a cut contained in I, then we can push forward the Loeb measure to a probability measure on C_U, which we also refer to as Loeb measure.

This procedure does in fact agree with Lebesgue measure in the case of the cuts U_N:

Theorem 9. *Given an infinite N, the pushforward via π_N of the Loeb measure on $[0, N)$ is the Lebesgue measure on $[0, 1]$.*

Motivated by the preceding discussion, Keisler and Leth asked whether the analog of Steinhaus' result holds for arbitrary cut spaces. Renling Jin answered this question affirmatively:

Theorem 10 ([14]). *Suppose that $I = [0, N)$ is an infinite interval and U is a cut contained in I. If A and B are definable subsets of I with positive Loeb measure, then $\pi_U(A + B)$ contains an interval.*

Technically speaking, in order for this to literally be true, one needs to assume that A and B are contained in the first half of I; otherwise, one needs to view I as a nonstandard cyclic group and then look at $A + B$ in the group sense. Without going into too much detail, the theorem is proven by contradiction, taking a "maximal" counterexample, and performing some nontrivial nonstandard counting.

Finishing the proof of Jin's theorem. We now have all of the pieces necessary to prove the sumset theorem. Suppose that A and B are subsets of $\mathbb{N}$ with positive Banach density. By the nonstandard characterization of limit, there is some infinite interval I such that $\mathrm{BD}(A)$ is approximately equal to $\frac{|A^* \cap I|}{|I|}$. In other words, $\mathrm{BD}(A) = \mu_I(A^* \cap I)$. Likewise, there is an infinite interval J such that $\mathrm{BD}(B) = \mu_J(B^* \cap J)$. Without loss of generality, we may assume that $|I| = |J| = N$ for some infinite N. Let a and b denote the left endpoints of I and J respectively and set $C := (A^* \cap I) - a$ and $D := (B^* \cap J) - b$. Note then that C and D are definable subsets of $[0, N)$ of positive Loeb measure. Without loss of generality, we may assume that C and D belong to the first half of $[0, N)$. Then by Jin's theorem mentioned above, $\pi_{\mathbb{N}}(C + D)$ contains an interval in $C_{\mathbb{N}}$. Translating back by $a + b$, one finds an infinite hyperfinite interval on which $A^* + B^*$ has only finite gaps, whence, by the nonstandard characterization described earlier, we see that $A + B$ is piecewise syndetic.

The Ultraproduct Construction

Of course, the lingering question remains: given a structure $\mathbf{S}$, is there a structure $\mathbf{S}^*$ satisfying (NA1)-(NA3)? Model theorists know these axioms to be consistent by basic model-theoretic facts. However, in this section, we present a construction that is much more "mainstream" and easy to describe to nonlogicians.

Obtaining $\mathbf{R}^*$ as an ultrapower. As in the passage from any number system to an extension where we are trying to add desired elements (e.g., the passage from $\mathbf{N}$ to $\mathbf{Z}$ to $\mathbf{Q}$ to $\mathbf{R}$ to $\mathbf{C}$), we simply formally add the new desired elements and then see what technicalities we need to introduce to make this formal passage precise. In this case, in passing from $\mathbf{R}$ to $\mathbf{R}^*$, we are trying to add infinite elements. Following the Cauchy sequence construction of $\mathbf{R}$ from $\mathbf{Q}$, we can simply add the sequence $1, 2, 3, \ldots$ to $\mathbf{R}$ and view this sequence as an infinite element of $\mathbf{R}^*$. Of course, just as in the case of the passage from $\mathbf{Q}$ to $\mathbf{R}$, many sequences should represent the same element of $\mathbf{R}^*$, e.g., the sequence $-32, \pi, 46, 4, 5, 6 \ldots$ should represent the same sequence as $1, 2, 3, \ldots$. In general, we should identify two sequences if they agree on a *big* number of indices, where two sequences (x_n) and (y_n) agree on a big number of terms if the set of n for which $x_n = y_n$ is a *large* subset of $\mathbf{N}$. Admittedly, the words "big" and "large" are rather vague here, so we need to isolate some properties that large subsets of $\mathbf{N}$ should have; the resulting notion is that of a *filter* on $\mathbf{N}$. Since the definition makes perfect sense for an arbitrary index set $\mathcal{I}$, we do so:

Definition 11. Suppose that $\mathcal{I}$ is a set. A (proper) *filter* on $\mathcal{I}$ is a collection $\mathcal{F}$ of subsets of $\mathcal{I}$ satisfying:

- $\emptyset \notin \mathcal{F}, \mathcal{I} \in \mathcal{F}$;
- if $A \in \mathcal{F}$ and $A \subseteq B \subseteq \mathcal{I}$, then $B \in \mathcal{F}$;
- if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

Denoting the set of $\mathcal{I}$-sequences by $\mathbf{R}^{\mathcal{I}}$, the above definition of a filter was engineered so that the relation $\sim_{\mathcal{F}}$ on $\mathbf{R}^{\mathcal{I}}$ defined by $x \sim_{\mathcal{F}} y$ if and only if $\{i \in \mathcal{I} : x_i = y_i\} \in \mathcal{F}$ is an equivalence relation. We denote the equivalence class of an $\mathcal{I}$-sequence x from $\mathbf{R}^{I}$ by $[x]$, and the set of all such equivalence classes by $\mathbf{R}^{\mathcal{F}}$. By identifying an element $r \in \mathbf{R}$ with the constant sequence $c_r = (r, r, r, \ldots)$, we get a natural inclusion of $\mathbf{R}$ into $\mathbf{R}^{\mathcal{I}}$, and by passing to the equivalence class we obtain the natural inclusion of $\mathbf{R}$ into $\mathbf{R}^{\mathcal{F}}$. In a diagram:

$$\mathbf{R} \xrightarrow{r \mapsto c_r} \mathbf{R}^{\mathcal{I}} \xrightarrow{x \mapsto [x]} \mathbf{R}^{\mathcal{F}}$$

Note also that $\mathbf{R}^{\mathcal{F}}$ has a natural field structure on it extending the field structure on $\mathbf{R}$.

There are now two problems with leaving things at this level of generality. The first can be motivated by the desire to turn $\mathbf{R}^{\mathcal{F}}$ into an *ordered* field. The natural thing to do would be to declare $[x] < [y]$ if and only if $\{i \in \mathcal{I} : x_i < y_i\} \in \mathcal{F}$. However, it is entirely possible that, under the above definition, we may have $\sim_{\mathcal{F}}$-inequivalent sequences x and y for which $[x] \not< [y]$ and $[y] \not< [x]$. We can remedy this by adding one further requirement:

Definition 12. A filter $\mathcal{U}$ on $\mathcal{I}$ is called an *ultrafilter* if, for every $A \subseteq \mathcal{I}$, we have that either $A \in \mathcal{U}$ or $\mathcal{I} \backslash A \in \mathcal{U}$ (but not both).

Here is the other problem: suppose that $\mathcal{U}_{37}$ is the collection of subsets of $\mathbf{N}$ defined by declaring $A \in \mathcal{U}_{37}$ if and only if $37 \in A$. Although this hardly matches the intuition of gathering *large* subsets of $\mathbf{N}$, it is easy to see that $\mathcal{U}_{37}$ is in fact an ultrafilter on $\mathbf{N}$, called the *principal ultrafilter on* $\mathbf{N}$ *generated by* 37. Now notice that in $\mathbf{R}^{\mathcal{U}_{37}}$, every sequence is equivalent to the real number given by its 37th entry. In other words, the infinite element $1, 2, 3, \ldots$ that we tried to add is not infinite at all, but rather is identified with the very finite number 37. To avoid this triviality, we add the following requirement:

Definition 13. If $\mathcal{I}$ is an index set and $i \in I$, we call $\mathcal{U}_i := \{A \subseteq I \ : \ i \in A\}$ the *principal ultrafilter on* I *generated by* i. An ultrafilter $\mathcal{U}$ on I is called *principal* if it is of the form $\mathcal{U}_i$ for some $i \in I$ and is otherwise called *nonprincipal*.

We may now summarize: for any nonprincipal ultrafilter $\mathcal{U}$ on any index set $\mathcal{I}$, $\mathbf{R}^{\mathcal{U}}$ is a proper ordered field extension of $\mathbf{R}$. Now, the ultrafilter on a set I may also be viewed as a finitely additive $\{0, 1\}$-valued measure on $\mathcal{I}$, which, moreover, gives points measure 0 precisely when the ultrafilter is nonprincipal. In this way, the ultrapower construction of $\mathbf{R}$ bears much resemblance to the practice common in measure theory of identifying measurable functions if they agree almost everywhere.

Ultrapowers of structures. The approach of the previous subsection works to obtain, given an initial structure $\mathbf{S}$, structures $\mathbf{S}^*$ satisfying (NA1) and (NA2). Indeed, given any basic set S_i, we can set $S_i^* := S_i^{\mathcal{U}}$ and given any basic relation R_j, we can set $R_j^* := R_j^{\mathcal{U}}$ and $\mathbf{S}^* = \mathbf{S}^{\mathcal{U}}$. In this case, $\mathbf{S}^*$ is called the *ultrapower* rather than an ultraproduct. As before, we have the natural inclusions, and it is clear that $\mathbf{S}^*$ satisfies (NA1). The fact that (NA2) holds in this context is an instance of a result in model theory known as *Łos' theorem*, which is easily proven by induction on the "complexity" of definable sets. Further, that this construction works for any initial structure $\mathbf{S}$ is the fact that one invokes to show that certain axiomatic theories of nonstandard methods are conservative over certain axiomatic theories of real numbers: i.e., anything that the former theory proves the latter theory also proves (cf. [5]).

Getting richer. It turns out that obtaining countable richness via ultrapowers is quite straightforward:

Theorem 14 (Keisler [15]). *Suppose that* $\mathcal{U}$ *is a nonprincipal ultrafilter on a countably infinite index set* $\mathcal{I}$. *Then for any structure* $\mathbf{S}$, $\mathbf{S}^{\mathcal{U}}$ *is countably rich.*

What about higher richness? If one insists on only using ultraproducts as a means of producing nonstandard extensions, then one can indeed obtain nonstandard universes with higher richness properties at the expense of dealing with some messy infinite combinatorics. Indeed, Keisler isolated a combinatorial property of ultrafilters, called *goodness*, and proved the following theorem:

Theorem 15 (Keisler [16], [17]). *Let* $\mathcal{U}$ *be an ultrafilter on a set* $\mathcal{I}$. *Then* $\mathcal{U}$ *is good if and only if: for every structure* $\mathbf{S}$, $\mathbf{S}^{\mathcal{U}}$ *is* maximally rich.

Here, we are using the admittedly vague term *maximally rich* to mean that the structure is as rich as the cardinality of its underlying domain allows. In order for this theorem to be useful, one needs to know that good ultrafilters exist. This is indeed the case: Keisler first proved, under the assumption of the Generalized Continuum Hypothesis, that any infinite set possesses a good ultrafilter; later, Kunen proved this fact without any extra set-theoretic assumption.

Other Approaches

Ever since its inception, a slew of different frameworks for approaching nonstandard analysis have been presented. Many of these approaches can be viewed as attempts to axiomatize different aspects of the ultraproduct construction. In this section, we briefly describe the differences between these approaches and our preferred approach using (NA1)-(NA3).

One alternative approach is just to axiomatize the behavior of the embedding of the original structure in its nonstandard extension. Depending on how formal one sought to be, one might then replace informal descriptions of the original structure by some axiomatic characterization of it. However, as the ultraproduct construction itself shows, first-order axioms cannot describe infinite structures up to isomorphism. Hence, if one sought first-order axiomatic characterizations, one might seek to add on axioms suggestive of the way in which the original structure was more canonical than its nonstandard extension. For instance, in the case of reals, one might add on first-order axioms to the effect that any bounded subset of $\mathbf{R}$ first-order definable by recourse to $\mathbf{R}^*$ had a least-upper bound (e.g., Nelson's [19, 1166] principle of standardization). Similar expressive difficulties emerge when one tries to find first-order axiomatic renditions of the richness conditions (e.g., Nelson's [19, 1166] principle of idealization, and Hrbacek's principle of bounded idealization). Since (NA1) -(NA3) makes no pretense to be a first-order axiomatization, it can avoid these niceties and just be content with our informal understanding of the real numbers and the richness conditions. In this, (NA1)-(NA3) is similar to the superstructure approach of Robinson–Zakon (cf. the Chang–Keisler model theory book), which additionally builds in the Mostowski collapse mentioned in "Power Sets and Internal Sets" and adopts a formulation of richness which does not require the notion of definable set.

A particularly influential axiomatic approach was that of Nelson's internal set theory ([19]). Nelson's aim was to create an overall set theory for nonstandard methods. In addition to the usual set-theoretic axioms, it also contained the aforementioned principles of standardization and idealization. Nelson's approach seems natural given that much of mathematics can be replicated in set theory. However, as we have sought to stress in the applications in "Selected Classical and Recent Applications," many of the most exciting applications of nonstandard analysis are to local areas of mathematics. Hence, as far as applications go, little seems to be gained by adopting the very global perspective of set theory.

Another recent approach pioneered by Benci and di Nasso [1] is to axiomatize the inclusion of the set of $\mathbf{R}$-valued sequences into the ultraproduct $\mathbf{R}^*$. Their theory is called *α-theory* because they axiomatize the map $x \mapsto [x]$ sending a *sequence* to its equivalence class (cf. the diagram after Definition 11). They use the notation $x \mapsto x[\alpha]$, since this notation reminds one of field extensions. For instance, it follows from the ultraproduct construction that the operation sending x to $x[\alpha]$ commutes with addition and multiplication:

$$x[\alpha]+y[\alpha] = (x+y)[\alpha], \qquad x[\alpha]\cdot y[\alpha] = (x\cdot y)[\alpha]$$

The idea of α-theory is to take these identities—and others pertaining to sets of reals—as axioms, and to derive transfer from these. The choice between α-theory and (NA1)-(NA3) is similar then to the choice between Dedekind cuts and Cauchy sequences: they are both equally good descriptions of their subject matter, but they just differ in which basic properties are derived and which are taken as primitive. One reason traditionally given for preferring Cauchy sequences over Dedekind cuts is that it more readily generalizes to other situations, e.g., complete metric spaces. Much the same is true of (NA1)-(NA3): since we can take extensions of any structure, the approach works just the same for, e.g., nonstandard p-adic analysis as for nonstandard real-analysis. By contrast, with α-theory, in each case one has to isolate some basic axioms that suffice for the derivation of the full transfer principle, and in each case one has to reprove the full transfer principle.

Another recent approach due to Benci and di Nasso [2] seeks to characterize the nonstandard extensions of the reals as images of certain rings (they also have similar results for certain classes of spaces). While they state their ring-theoretic results for the reals, their proof generalizes as follows (recall the notion of complete structure given immediately after Definition 6):

Theorem 16. *Suppose that $\mathbf{K}$ is the complete structure of an uncountable ordered field. Then $\mathbf{K}^*$ satisfies (NA1)-(NA2) if and only if $\mathbf{K}^*$ is the image under a ring homomorphism of a composable ring over $\mathbf{K}$.*

In this, the ring is said to be *composable over* $\mathbf{K}$ if it is a subring of the ring $\mathbf{K}^{\mathcal{I}}$ of all functions from some index set $\mathcal{I}$ to the original field $\mathbf{K}$ which is closed under taking compositions with functions from $\mathbf{K}$ to $\mathbf{K}$ (and which contains all the constant functions). However, this of course is not to say that studying composable rings is always the best way to study nonstandard methods. After all, any group is the image of a free group, but it would be a mistake to approach all problems in group theory through the lens of free groups.

A unifying motivation behind these recent approaches is the desire for a more mathematically natural framework for nonstandard analysis. By contrast, Robinson himself was an instrumentalist about the methods he developed (cf. [5]): He thought that their worth was tied less to any notions of naturalness and more to their proven track record in obtaining results about standard structures, such as we have surveyed in "Selected Classical and Recent Applications." As we have sought to emphasize throughout, nonstandard methods such as (NA1)-(NA3) are easy to state and use, and their consistency is easily verifiable via the ultraproduct construction. In describing the ultraproduct construction we mentioned the analogy with the Lebesgue integral. Here is another respect in which they are similar: to use the Lebesgue integral correctly, one does not need to keep its measure-theoretic construction constantly in view, but rather one can just work with characteristic properties of it, like the Dominated Convergence Theorem. Likewise, to use nonstandard methods correctly, one does not need to keep the ultraproduct construction constantly in mind. Rather, as we sought to illustrate with Jin's Theorem in "The Axioms in Action: Jin's Theorem" one can do nonstandard analysis just using the three simple axioms (NA1)-(NA3).

References

[1] Benci V, Di Nasso M. Alpha-theory: an elementary axiomatics for nonstandard analysis, *Expo. Math.*, no. 4 (21):355–386, 2003, DOI 10.1016/S0723-0869(03)80038-5. MR2022004

[2] Benci V, Di Nasso M. A purely algebraic characterization of the hyperreal numbers, *Proc. Amer. Math. Soc.*, no. 9 (133):2501–2505, 2005, DOI 10.1090/S0002-9939-05-07429-0. MR2146232

[3] Bernstein AR, Robinson A. Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos, *Pacific J. Math.* (16):421–431, 1966. MR0193504

[4] Breuillard E, Green B, Tao T. The structure of approximate groups, *Publ. Math. Inst. Hautes Études Sci.* (116):115–221, 2012, DOI 10.1007/s10240-012-0043-9. MR3090256

[5] Button T, Walsh S. *Philosophy and model theory. Chapter 4: Compactness, infinitesimals, and the reals*, Oxford University Press, Oxford, 2018. With a historical appendix by Wilfrid Hodges. MR3821510

[6] Di Nasso M, Goldbring I, Lupini M. *Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory*, Lecture Notes in Mathematics, Springer, 2019.

[7] van den Dries L, Goldbring I. Hilbert's 5th problem, *Enseign. Math.*, no. 1-2 (61):3–43, 2015, DOI 10.4171/LEM/61-1/2-2. MR3449281

[8] van den Dries L, Wilkie A. Gromov's theorem on groups of polynomial growth and elementary logic, *J. Algebra*, no. 2 (89):349–374, 1984, DOI 10.1016/0021-8693(84)90223-0. MR751150

[9] Goldblatt R. *Lectures on the hyperreals*, Graduate Texts in Mathematics, vol. 188, Springer–Verlag, New York, 1998. An introduction to nonstandard analysis. MR1643950

[10] Goldbring I. Hilbert's fifth problem for local groups, *Ann. of Math. (2)*, no. 2 (172):1269–1314, 2010, DOI 10.4007/annals.2010.172.1273. MR2680491

[11] Gromov M. Groups of polynomial growth and expanding maps, *Inst. Hautes Études Sci. Publ. Math.* (53):53–73, 1981. MR623534

[12] Hirschfeld J. The nonstandard treatment of Hilbert's fifth problem, *Trans. Amer. Math. Soc.*, no. 1 (321):379–400, 1990, DOI 10.2307/2001608. MR967314

[13] Hrushovski E. Stable group theory and approximate subgroups, *J. Amer. Math. Soc.*, no. 1 (25):189–243, 2012, DOI 10.1090/S0894-0347-2011-00708-X. MR2833482

[14] Jin R. Nonstandard methods for upper Banach density problems, *J. Number Theory*, no. 1 (91):20–38, 2001, DOI 10.1006/jnth.2001.2671. MR1869316

[15] Keisler J. Ultraproducts and elementary classes, *Nederl. Akad. Wetensch. Proc. Ser. A 64 = Indag. Math.* (23):477–495, 1961. MR0140396

[16] Keisler J. Ultraproducts and saturated models, *Nederl. Akad. Wetensch. Proc. Ser. A 67 = Indag. Math.* (26):178–186, 1964. MR0168483

[17] Keisler J. Ultraproducts which are not saturated, *J. Symbolic Logic* (32):23–46, 1967, DOI 10.2307/2271240. MR0218224

[18] Mann A. *How groups grow*, London Mathematical Society Lecture Note Series, vol. 395, Cambridge University Press, Cambridge, 2012. MR2894945

[19] Nelson E. Internal set theory: a new approach to nonstandard analysis, *Bull. Amer. Math. Soc.*, no. 6 (83):1165–1198, 1977, DOI 10.1090/S0002-9904-1977-14398-X. MR0469763

[20] Robinson A. *nonstandard analysis*, North-Holland Publishing Co., Amsterdam, 1966. MR0205854

Isaac Goldbring

Sean Walsh

Credits

Opening image is courtesy of Getty Images, by Pobytov, from the DigitalVision Vectors Collection.

Author photos are courtesy of Isaac Goldbring and Sean Walsh.