



Hyperreals and Non-Standard Analysis

GSO Talk: Isaac Solomon

1. HISTORY: NEWTON, LEIBNIZ, EULER, BERKELEY AND ROBINSON

With the introduction of the differential and integral calculus in the latter end of the seventeenth century, mathematicians had begun to grapple with the notion of the infinite in a very direct way. Lacking the $\epsilon - \delta$ definition of the limit, they had no formal way of expressing quantities that were arbitrarily large or small. Instead, they tentatively embraced the infinitesimal: a non-zero entity that was yet smaller than any finite number.

Leibniz made use of the infinitely small quantities dx, dy in his development of the calculus. His actual view on the legitimacy of infinitesimals can be seen in the following quote:

“It will be sufficient if, when we speak of infinitely great (or more strictly unlimited), or of infinitely small quantities (i.e., the very least of those within our knowledge), it is understood that we mean quantities that are indefinitely great or indefinitely small, i.e., as great as you please, or as small as you please, so that the error that any one may assign may be less than a certain assigned quantity ... If any one wishes to understand these (the infinitely great or infinitely small) as the ultimate things, or as truly infinite, it can be done, and that too without falling back upon a controversy about the reality of extensions, or of infinite continuums in general, or of the infinitely small, say, even though he think that such things are impossible; it will be sufficient simply to make use of them as a tool that has advantages for the purpose of the calculation, just as the algebraists retain imaginary roots with great profit. For they contain a handy mean of reckoning, as can manifestly be verified in every case in a rigorous manner by the method already stated.”

Newton related fluxions \dot{x}, \dot{y} to their respective fluents x, y as geometric entities:

“the speeds with which they flow and are increased by their generating motion.”

In his later writings, Newton attempted to explain his theory of fluxions in terms of limits of ratios of quantities, a notion not dissimilar to that of the modern limit.

Euler displayed a regrettably haphazard attitude in the use of infinitesimals. He simply assumed they existed, and used them like finite numbers. The following excerpt from his “Introduction to the Analysis of the Infinite” (1748) characterizes his approach.

“Let ω be an infinitely small number, or a fraction so small that, although not equal to zero, still $a^\omega = 1 + \psi$, where ψ is also an infinitely small number...”

The Anglo-Irish philosopher George Berkeley criticized the foundations of Newton's calculus.

“And what are these fluxions? ... May we not call them the ghosts of departed quantities?”

Attitudes like this led to infinitesimals being discarded as the basis for analysis, being replaced by the rigorous notion of the limit developed by Cauchy, Weierstrass, and others. It was not until three hundred years later that Abraham Robinson, then a professor at UCLA, reintroduced infinitesimals. In his 1966 text “Non-Standard Analysis” he writes:

“In the fall of 1960 it occurred to me that the concepts and methods of contemporary Mathematical Logic are capable of providing a suitable framework for the development of the Differential and Integral Calculus by means of infinitely small and infinitely large numbers.”

In this short article, I intend to explore some of the central results and ideas in the theory of Hyperreal Numbers and Non-Standard Analysis. To conclude, we will give applications to topology and history.

Texts Used (Needs Proper Bibliography): Lectures on the Hyperreals: An Introduction to Nonstandard Analysis. Robert Goldblatt. Springer.

Herbert, Enderton: A Mathematical Introduction to Logic...etc.

2. CONSTRUCTIONS OF ${}^*\mathbb{R}$

2.1. The Compactness Theorem. Underlying Non-Standard Analysis is an enlarged field of real numbers, denoted ${}^*\mathbb{R}$, that contains infinitesimal-like elements. This structure is called the Hyperreals, and we shall construct it in two ways. The first is purely model-theoretic, and the second is based on techniques of logic and algebra.

Compactness Theorem. *Any set of first-order sentences has a model if and only if every finite subset of it has a model.*

Let $\text{Th } \mathbb{R}$ be the theory of the real field: the collection of all first-order sentences in the signature of \mathbb{R} that are satisfied by \mathbb{R} . Let ψ_n be the sentence $x > \overbrace{S(S(S(\dots(0))))}^n$ for $n \in \mathbb{N}$. Lastly, let

$$\Gamma = \text{Th } \mathbb{R} \cup \left[\bigcup_{n \in \mathbb{N}} \psi_n \right]$$

Any finite subset of Γ has a model: \mathbb{R} itself. This can be seen by assigning x to some quantity larger than the maximum of the set $\{n_1, \dots, n_k\}$ represented in the finite collection of sentences. By the Compactness Theorem, Γ also has a model. However, because \mathbb{R} contains no largest element, it cannot be a model for Γ : there is no element to which x can be assigned. Instead, we call our model ${}^*\mathbb{R}$, and the element x an “unlimited element of ${}^*\mathbb{R}$.” Symmetrically, $1/x$ is an infinitesimal in ${}^*\mathbb{R}$.

The problem with this construction is that is non-constructive. We simply invoked the Compactness Theorem and were handed a non-standard model of analysis. As a result, the elements of ${}^*\mathbb{R}$ are somewhat intangible. To rectify this situation, we abandon model theory for set theory, in the form of an Ultrapower.

2.2. Equivalence Classes of Cauchy Sequences. In Cantor’s construction of the reals, he projects the set of rational Cauchy Sequences, $\mathcal{C}_{\mathbb{Q}}$ onto the real line:

$$\mathcal{C}_{\mathbb{Q}} \rightarrow \mathbb{R}$$

It has to be said, however, that this projection is rather coarse, for it groups together all sequences within a common point of convergence, regardless of the way in which they converge. What would get if we imposed a finer equivalence relation on the set of sequences? For example: what if we considered two sequences equivalent if their points agreed on a sufficiently large set? We shall see that this will give rise to ${}^*\mathbb{R}$.

First and foremost, we need to make precise the notion of a “sufficiently large” set. By considering those properties that we would expect our collection of large sets to have, it will become apparent exactly how to define them.

- (1) If A is a large set, and $A \subset B$, then B should also be large.
- (2) If a_n is equivalent to b_n , and b_n is equivalent to c_n , we would like a_n to be equivalent to c_n . Thus if A, B are large sets, $A \cap B$ should also be large.

- (3) \mathbb{N} should be large, so that a sequence is equivalent to itself. Conversely, \emptyset should not be large, otherwise every set would be large by (1).
- (4) Let $a_n = 1111111\dots$, $b_n = 0000000\dots$, and $c_n = 0101010\dots$. Clearly, a_n and b_n should not be equivalent. However, if both the even and odd naturals are large, c_n will be equivalent to both a_n and b_n , which will be equivalent by transitivity. Thus we stipulate that A and $\mathbb{N} \setminus A$ cannot both be large.

2.3. Filters and Ultrafilters. Let I be a nonempty set. A **filter** on the power set of I is a nonempty collection $\mathcal{F} \subset \mathcal{P}(I)$ that satisfies the following two axioms:

- If $A \in \mathcal{F}$, and $A \subset B \subset I$, then $B \in \mathcal{F}$.
- If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

A filter is **proper** if it does not contain \emptyset . Lastly, an **ultrafilter** is a proper filter with the additional axiom that

- For all $A \subset I$, either $A \in \mathcal{F}$ or $I \setminus A \in \mathcal{F}$.

In other words, an ultrafilter has an opinion on every subset of I , whether it is in the filter, or whether its complement is in the filter. Some examples of filters are:

- (1) The **cofinite filter**, $\mathcal{F}^{co} = \{A \subset I : I \setminus A \text{ is finite}\}$.
- (2) Let $\emptyset \neq \mathcal{H} \subset \mathcal{P}(I)$. The **filter generated by \mathcal{H}** is

$$\mathcal{F}^{\mathcal{H}} = \{A \subset I : A \supset B_1 \cap \dots \cap B_n\}$$

for some $B_i \in \mathcal{H}$. This is the smallest filter on I containing \mathcal{H} .

Lastly, a filter is **principal** if it contains a one-element set. Note that if \mathcal{F} is a non-principal ultrafilter, it cannot contain a finite set. For if it did, we could use the fact that filters are closed under intersections to show that \mathcal{F} must contain a one-element set, and hence would be principal. Thus \mathcal{F} must contain all cofinite sets.

At this point, it is apparent that the notion of a collection of large sets is captured perfectly in our definition of a non-principal ultrafilter on \mathbb{N} .

We now prove that such ultrafilters exist.

Theorem. *Make $\mathcal{P}(I)$ a poset ordered by set-inclusion. Then a filter is an ultrafilter if and only if it is a maximal proper filter.*

Proof. Suppose that \mathcal{F} is an ultrafilter, but is not a maximal proper filter. Then extend \mathcal{F} to \mathcal{F}' , a maximal proper filter. Hence there is a set $S \in \mathcal{F}'$ such that $S \notin \mathcal{F}$. So $I \setminus S \in \mathcal{F}$, since it is an ultrafilter. Thus $I \setminus S \in \mathcal{F}'$ and

$$(I \setminus S) \cap (S) = \emptyset \in \mathcal{F}'$$

and then \mathcal{F}' wouldn't be proper.

Conversely, suppose \mathcal{F} is a maximal, proper filter that is not an ultrafilter. Then, there exists a set A such that $A \notin \mathcal{F}$ and $I \setminus A \notin \mathcal{F}$. Now, if there exists a set $B \in \mathcal{F}$ such that $B \cap A = \emptyset$, then $I \setminus A \in \mathcal{F}$, since it is maximal, contradicting our hypothesis. Otherwise, $\mathcal{F} \cup A$ can be extended to a proper filter, contradicting our hypothesis that \mathcal{F} is maximal. \square

Theorem. *A collection $\mathcal{H} \subset \mathcal{P}(I)$ is said to have the finite intersection property if the intersection of every finite, nonempty subcollection of \mathcal{H} is nonempty. If \mathcal{H} has the finite intersection property, then $\mathcal{F}^{\mathcal{H}}$ can be extended to an ultrafilter on I .*

Proof. If \mathcal{H} has the finite intersection property, then $\mathcal{F}^{\mathcal{H}}$ must be proper. Let P be the collection of all proper filters on I that include $\mathcal{F}^{\mathcal{H}}$, partially ordered by set inclusion. Then every linearly ordered subset of P has an upper bound in P , since the union of this chain is also a filter in P . By Zorn's Lemma, P has a maximal element, denoted \mathcal{G} . By the previous theorem, \mathcal{G} is an ultrafilter. \square

Corollary. *Any infinite set has a nonprincipal ultrafilter on it.*

Proof. If I is infinite, then \mathcal{F}^{co} is proper and has the finite intersection property, and so can be extended to an ultrafilter on I , denoted \mathcal{F} . However, \mathcal{F} cannot be a principal ultrafilter, since for all point-sets $\{i\}$, $I \setminus \{i\} \in \mathcal{F}^{co} \subset \mathcal{F}$. \square

2.4. Ultrapower Construction. Let $\mathbb{R}^{\mathbb{N}}$ be the set of all real-valued sequences. Define addition and multiplication on $\mathbb{R}^{\mathbb{N}}$ by

$$r_n + s_n = (r + s)_n, \quad r_n \cdot s_n = (r \cdot s)_n$$

Then $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ is a commutative ring with zero $(0, 0, 0, \dots)$ and identity $(1, 1, 1, \dots)$, and is example of what is known as a **direct power**. Let \mathcal{F} be a nonprincipal ultrafilter on \mathbb{N} , and define the equivalence relation $\sim_{\mathcal{F}}$ as follows:

$$r_n \sim_{\mathcal{F}} s_n := \{k \in \mathbb{N} : r_k = s_k\} \in \mathcal{F}$$

In other words, two sequences are equivalent if the set of naturals on which their points agree is in our nonprincipal ultrafilter. In future, we will write $[[r_n = s_n]]$ in place of $\{k \in \mathbb{N} : r_k = s_k\}$. Finally, define

$${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}} / \sim_{\mathcal{F}}$$

The quotient space of a direct product defined by congruence with respect to an ultrafilter is called a **ultrapower**, and this is precisely how we have defined ${}^*\mathbb{R}$. Similarly, we could have defined the hypernaturals ${}^*\mathbb{N}$, hyperintegers ${}^*\mathbb{Z}$, and hyperrationals ${}^*\mathbb{Q}$. The elements of these systems are congruence classes of sequences, whose members are written $[r_n]$, or $[r]$ for short. To be certain that this definition of ${}^*\mathbb{R}$ is useful, we prove

Theorem. *The structure $({}^*\mathbb{R}, +, \cdot, <)$ is an ordered field with zero $[(0, 0, 0\dots)]$ and and unity $[(1, 1, 1\dots)]$.*

Proof. Trivial. \square

To regain our bearings, we identify each real number r with the constant sequence (r, r, r, \dots) . The element $\omega = (1, 2, 3, \dots)$ would be infinitely large, or unlimited, and its inverse, $\epsilon = (1, 1/2, 1/3, \dots)$ would be an infinitesimal. In fact, every sequence converging to zero that is not eventually constant becomes an infinitesimal, and every diverging sequence becomes an unlimited number. This sounds like a lot of extra numbers. However,

Theorem. \mathbb{R} and ${}^*\mathbb{R}$ have the same cardinality.

Proof. $f(r) \rightarrow [(r, r, r, \dots)]$ is an injection from \mathbb{R} to ${}^*\mathbb{R}$. Conversely, we know that \mathbb{R} and $\mathbb{R}^{\mathbb{N}}$ have the same cardinality, so let f be an injective function from $\mathbb{R}^{\mathbb{N}}$ to \mathbb{R} . Using choice, pick a representative sequence r_n from each equivalence class, and map that equivalence class to $f(r_n)$. \square

2.5. Extending Sets, Functions and Relations. For $A \subset \mathbb{R}$, we define

$$[r] \in {}^*A \iff [[r_n \in A]] \in \mathcal{F}$$

If $f : A \rightarrow \mathbb{R}$, with $A \subset \mathbb{R}$, we can extend f to a function ${}^*f : {}^*A \rightarrow {}^*\mathbb{R}$ by defining

$$s_n = \begin{cases} f(r_n) & \text{if } n \in [[r_n \in A]] \\ 0 & \text{otherwise} \end{cases}$$

and setting ${}^*f([r]) = [s]$.

Lastly, if P is a k -ary relation on \mathbb{R} , define *P

$${}^*P([r^1], \dots, [r^k]) \iff [[P(r_n^1, \dots, r_n^k)]] \in \mathcal{F}$$

3. UNIQUENESS OF ${}^*\mathbb{R}$

It turns out that if you assume the Continuum Hypothesis (CH), the choice of nonprincipal ultrafilter is irrelevant. You will get a unique structure, up to isomorphism. Unfortunately, without CH, this result does not hold.

The proof of the first claim is involved, and is based on a result by Erdős, Gillman and Henriksen in 1955. For a proof of the second result, read the paper “Non-isomorphic hyper-real fields from non-isomorphic ultrapowers,” *Mathematische Zeitschrift*, 181:93 - 96, 1982 by J. Roitman. I do not know if the UCLA library has this journal on hold.

4. LOS TRANSFER PRINCIPLE

Also in 1955, the Polish mathematician Jerzy Łoś proved the following (somewhat) surprising result. Any first-order statement about \mathbb{R} is also true about ${}^*\mathbb{R}$, and visa versa, if properly translated. Moreover, by the Completeness Theorem, what can be proven in one can be proven in the other.

Theorem. *Let $\Phi(X_1, \dots, X_n, x_1, \dots, x_n)$ be a formula in the first order logic of \mathbb{R} . Then for any $A_1, \dots, A_m \subseteq \mathbb{R}$ and ${}^*r_1, \dots, {}^*r_n \in \mathbb{R}$,*

$$\{i \in \mathbb{N} \mid \Phi(A_1, \dots, A_m, r_n^1, \dots, r_n^i) \in \mathcal{F}\}$$

iff

$$\Phi({}^*A_1, \dots, {}^*A_m, {}^*r_1, \dots, {}^*r_n) \text{ is true in } {}^*\mathbb{R}.$$

There is no subtle and wonderful trick at play here, and this theorem needs to be verified via case by case induction. Here is a sample case: suppose that we have two formulae Φ_1 and Φ_2 about \mathbb{R} for which the theorem holds. Let $\Phi = \Phi_1 \wedge \Phi_2$. The result holds for Φ by the finite intersection property.

Now, we can prove the Transfer Principle

Theorem. *Let $\Phi(X_1, \dots, X_n, x_1, \dots, x_n)$ be a formula in the first order logic of \mathbb{R} . Then for any $A_1, \dots, A_m \subseteq \mathbb{R}$ and $r_1, \dots, r_n \in \mathbb{R}$, $\Phi(A_1, \dots, A_m, r_1, \dots, r_n)$ is true in \mathbb{R} iff $\Phi({}^*A_1, \dots, {}^*A_m, {}^*r_1, \dots, {}^*r_n)$ is true in ${}^*\mathbb{R}$.*

Proof. From the above theorem, we see that

$$\{i \in \mathbb{N} \mid \Phi(A_1, \dots, A_m, r_n^1, \dots, r_n^i) \in \mathcal{F}\}$$

iff

$$\Phi({}^*A_1, \dots, {}^*A_m, {}^*r_1, \dots, {}^*r_n) \text{ is true in } {}^*\mathbb{R}.$$

But the set $\{i \in \mathbb{N} \mid \Phi(A_1, \dots, A_m, r_n^1, \dots, r_n^i) \in \mathcal{F}\} = \mathbb{N}$ is in the ultrafilter if Φ is true of $A_1, \dots, A_m, r_1, \dots, r_n$, and is equal to $\emptyset \notin \mathcal{F}$ otherwise. \square

5. THE STRUCTURE OF ${}^*\mathbb{R}$

5.1. **Limited, Unlimited, Infinitesimal, and Appreciable Numbers.** Time for some definitions. Let $x \in {}^*\mathbb{R}$, and extend the absolute value function to ${}^*\mathbb{R}$ in the obvious way. Then

- x is **limited** if $|x| < n$ for some $n \in \mathbb{N}$.
- x is **unlimited** if $|x| > n$ for all $n \in \mathbb{N}$.
- x is **infinitesimal** if $|x| < 1/n$ for all $n \in \mathbb{N}$.
- x is **appreciable** if $1/n < |x| < n$ for some $n \in \mathbb{N}$.

Observe that the sets \mathbb{L} and \mathbb{I} of limited and infinitesimals are in fact subrings of ${}^*\mathbb{R}$.

5.2. **Halos, Galaxies, and Shadows.** Define the equivalence relation $\sim_{\mathbb{I}}$ on ${}^*\mathbb{R}$ as follows:

$$a \sim_{\mathbb{I}} b \iff a - b \in \mathbb{I}$$

Define the equivalence relation $\sim_{\mathbb{L}}$ on ${}^*\mathbb{R}$ as follows:

$$a \sim_{\mathbb{L}} b \iff a - b \in \mathbb{L}$$

The following naming conventions were popularized by the French school of nonstandard analysis.

- The **Halo** of an element $x \in {}^*\mathbb{R}$ is its equivalence class under $\sim_{\mathbb{I}}$. This is the set of all hyperreals “infinitely” close to x .
- The **Galaxy** of an element $x \in {}^*\mathbb{R}$ is its equivalence class under $\sim_{\mathbb{L}}$. This is the set of all hyperreals a limited distance away from x .

Theorem. *Every limited hyperreal x is infinitely close to a unique real number, called its **shadow**, and written $sh(x)$.*

Proof. Let $A = \{r \in \mathbb{R} \mid r < x\}$. Since x is limited, this set is non-empty, and bounded above. By the completeness of \mathbb{R} , it has a least upper bound c . We claim $c \sim_{\mathbb{I}} x$.

To see this, let $\epsilon > 0$ be any positive real number. We must have $x \leq c + \epsilon$, as c is a least upper bound. Hence $0 < x - c < \epsilon$. Letting $\epsilon \rightarrow 0$, we see that $x - c \in \mathbb{I}$.

Uniqueness follows from the fact that any two reals that are infinitely close must be equal. \square

5.3. $\mathbb{L}/\mathbb{I} \cong \mathbb{R}$. What happens if you consider the quotient ring \mathbb{L}/\mathbb{I} ? You are effectively collapsing all limited hyperreals together that are infinitely close. It should come as no surprise then that

Theorem. *The quotient ring \mathbb{L}/\mathbb{I} is isomorphic to the real number field \mathbb{R}*

Proof. The isomorphism is given by the correspondence $hal(x) \rightarrow sh(x)$. \square

Corollary. \mathbb{I} is a maximal ideal in \mathbb{L} .

6. NON-STANDARD ANALYSIS

6.1. **Limits.** A real-valued sequence $\{s_n : n \in \mathbb{N}\}$ is a function $\mathbb{N} \rightarrow \mathbb{R}$. We can extend f to

$${}^*f : {}^*\mathbb{N} \rightarrow \mathbb{R}$$

For $n \in {}^*\mathbb{N}_\infty$, we call s_n an **extended term** of the sequence, and $\{s_n : n \in {}^*\mathbb{N}_\infty\}$ the **extended tail** of a sequence.

Theorem. A real valued-sequence s_n convergents to $L \in \mathbb{R}$ iff $s_n \sim_{\mathbb{I}} L$ for all unlimited n .

Proof. Suppose that s_n converges to L . Then for any $\epsilon > 0$, there exists n_ϵ such that

$$(\forall n \in \mathbb{N})(n > n_\epsilon \rightarrow |s_n - L| < \epsilon)$$

By transfer,

$$(\forall n \in {}^*\mathbb{N})(n > n_\epsilon \rightarrow |s_n - L| < \epsilon)$$

But if N is unlimited, then $N > n_\epsilon$ for any ϵ , so $s_n \sim_{\mathbb{I}} L$.

Conversely, if $s_n \sim_{\mathbb{I}} L$ for all unlimited n , then take $\epsilon > 0$. We have

$$(\exists z \in {}^*\mathbb{N})(\forall n \in {}^*\mathbb{N})(n > z \rightarrow |s_n - L| < \epsilon)$$

By transfer,

$$(\exists z \in \mathbb{N})(\forall n \in \mathbb{N})(n > z \rightarrow |s_n - L| < \epsilon)$$

which is the very definition of convergent to L .

□

6.2. Continuity and Uniform Continuity.

Theorem. f is continuous at real c iff $f(x) \sim_{\mathbb{I}} f(c)$ when $x \sim_{\mathbb{I}} c$.

Proof. Suppose that $(x \sim_{\mathbb{I}} c) \rightarrow (f(x) \sim_{\mathbb{I}} f(c))$. Pick any real $\epsilon > 0$ you damn please. Then the following statement is true:

$$(\exists \delta \in {}^*\mathbb{R}^+)(\forall x \in {}^*\mathbb{R})(|x - c| < \delta \rightarrow |f(x) - f(c)| < \epsilon)$$

In particular, one can let $\delta \in \mathbb{I}^+$. Now, by transfer,

$$(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(|x - c| < \delta \rightarrow |f(x) - f(c)| < \epsilon)$$

which gives us continuity!

Conversely, suppose we have continuity. Let $\epsilon > 0$ be any positive real. Then there exists a positive real δ so that

$$(\forall x \in \mathbb{R})(|x - c| < \delta \rightarrow |f(x) - f(c)| < \epsilon)$$

is true. By transfer,

$$(\forall x \in {}^*\mathbb{R})(|x - c| < \delta \rightarrow |f(x) - f(c)| < \epsilon)$$

Now, if $x \sim_{\mathbb{I}} c$, then $|x - c|$ is less than any positive real, so the condition is verified. Then $|f(x) - f(c)| < \epsilon$ holds for all positive ϵ , so $f(x) \sim_{\mathbb{I}} f(c)$ \square

More generally,

Theorem. *f is continuous on the real interval A iff $\forall c \in A$, we have $f(x) = f(c)$, for all $x \in \text{halo}(c) \subseteq {}^*A$.*

In English, this states that a function is continuous on a real interval A if, for any real point $c \in A$,

$$f(\text{hal}(c)) \subseteq \text{hal}(f(c))$$

What would happen if we demanded that this hold for all points $c \in {}^*A$?

Theorem. *f is uniformly continuous on A iff $x \sim_{\mathbb{I}} y$ implies $f(x) \sim_{\mathbb{I}} f(y)$ for all hyperreals $x, y \in {}^*A$.*

Proof. Suppose that $(x \sim_{\mathbb{I}} c) \rightarrow (f(x) \sim_{\mathbb{I}} f(c))$. Pick any real $\epsilon > 0$ that your heart desires. Then the following statement is true:

$$(\exists \delta \in {}^*\mathbb{R}^+)(\forall x \in {}^*\mathbb{R})(\forall y \in {}^*\mathbb{R})(|x - y| < \delta \rightarrow |f(x) - f(y)| < \epsilon)$$

In particular, one can let $\delta \in \mathbb{I}^+$. Now, by transfer,

$$(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(|x - y| < \delta \rightarrow |f(x) - f(y)| < \epsilon)$$

which gives us uniform continuity. Conversely, suppose we have uniform continuity. Let $\epsilon > 0$ be any positive real. Then there exists a positive real δ so that

$$(\forall x \in \mathbb{R})(\forall y \in \mathbb{R})(|x - y| < \delta \rightarrow |f(x) - f(y)| < \epsilon)$$

is true. By transfer,

$$(\forall x \in {}^*\mathbb{R})(\forall y \in {}^*\mathbb{R})(|x - y| < \delta \rightarrow |f(x) - f(y)| < \epsilon)$$

Now, if $x \sim_{\mathbb{I}} c$, then $|x - c|$ is less than any positive real, so the condition is verified. Then $|f(x) - f(c)| < \epsilon$ holds for all positive ϵ , so $f(x) \sim_{\mathbb{I}} f(c)$ \square

This allows us to give the following slick proof:

Theorem. *If f is continuous on the closed interval $[a, b] \in \mathbb{R}$, then it is uniformly continuous.*

Proof. Take hyperreals $x, y \in {}^*[a, b]$ so that $x \sim_{\mathbb{I}} y$. Let $c = \text{sh}(x)$. It is easy to see that $c \in [a, b]$. Since f is continuous, $f(x) \sim_{\mathbb{I}} f(c)$. Identically, we could show that $f(y) \sim_{\mathbb{I}} f(c)$. By the transitivity of this equivalence relation, we have $f(x) \sim_{\mathbb{I}} f(y)$. \square

Remark. *Take note of this result. When we define compactness in nonstandard terms, it will be immediately apparent how to show that a continuous function on a compact set is uniformly continuous.*

6.3. **Cauchy's Vindication.** In 1821, Cauchy wrote

"If the different terms of the series

$$u_0 + u_1 + \cdots + u_n + u_{n+1} + \cdots$$

are functions of the same variable x , continuous with respect to this variable in the neighborhood of a particular value for which the series is convergent, the sum s of the series is also, in the neighborhood of this particular value, a continuous function of x ."

Many people believe that this statement of Cauchy is a mistake, but it is possible to justify it.

Theorem. $\{f_n : n \in \mathbb{N}\}$ converge uniformly to the function $f : A \rightarrow \mathbb{R}$ if and only if for each $x \in {}^*A$ and each unlimited $n \in {}^*\mathbb{N}$, $f_n(x) \sim f(x)$.

Proof. Exercise. □

If Cauchy meant to assert the continuity of u_n for infinitely large n , and meant "the neighborhood of a particular value" to include points infinitely close to the value, then the result is correct. Admittedly, it is a bit of stretch. But, hey, this is Cauchy we are talking about here.

6.4. Defining Differentiation and Integration. Using our earlier non-standard characterization of the limit, we assert that

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = L$$

iff

$$\frac{f(x+\epsilon) - f(x)}{\epsilon} \sim_{\mathbb{I}} L$$

for every infinitesimal ϵ .

One can also define integration in a non-standard context. Let

$$S_a^b(f, P) = \sum_{i=1}^n f(x_{i-1}) \Delta x_i$$

be the standard Riemann sum. Fixing a, b and f , one can regard $S_a^b(f, *)$ as a function on \mathbb{R}^+ . Naturally, this can be extended to a function on ${}^*\mathbb{R}^+$. Thus, if Δx is an infinitesimal, we get a partition of $[a, b]$ such that $x_{i+1} - x_i = \Delta x$. Naturally, this partition has infinitely many values. Using this notation, we can define

$$\int_a^b f(x) dx = \text{shadow}(S_a^b(f, \Delta x))$$

for any positive infinitesimal Δx . Again, we shall not prove the equivalence. It follows by standard non-standard methods, such as transfer and a redefinition of limits.

6.5. Compactness.

Theorem. (*Robinson's Compactness Criterion*) B is a compact subset of \mathbb{R} iff every member of *B is infinitely close to some member of B .

Proof. Suppose that Robinson's criterion fails, so that there is a hyperreal $b \in {}^*B$ whose shadow is not in B . Then for each $r \in b$, there exists real $\epsilon_r > 0$ such that $|b - r| > \epsilon_r$. Consider the open cover

$$\mathcal{O} = \bigcup_{r \in B} N(r; \epsilon_r)$$

Suppose that \mathcal{O} has a finite subcover

$$B \subseteq (r_1 - \epsilon_{r_1}, r_1 + \epsilon_{r_1}) \cup \cdots \cup (r_n - \epsilon_{r_n}, r_n + \epsilon_{r_n})$$

Then

$${}^*B \subseteq {}^*(r_1 - \epsilon_{r_1}, r_1 + \epsilon_{r_1}) \cup \cdots \cup {}^*(r_n - \epsilon_{r_n}, r_n + \epsilon_{r_n})$$

which would imply $|b^* - r_i| < \epsilon_{r_i}$, contrary to our construction. Thus no finite subcover of \mathcal{O} exists, and B is not compact.

Conversely, suppose that B is not compact, and let $\mathcal{C} = \{A_i : i \in I\}$ be a cover of B with no finite subcover. For each $r \in B$, there exists rationals p_r, q_r such that

$$b \in (p_r, q_r) \subseteq A_i$$

The union of these rational-endpoints intervals is another open cover \mathcal{C}' of B . Like \mathcal{C} , \mathcal{C}' has no finite subcover. Moreover, \mathcal{C}' contains only countably many intervals. Thus we can write

$$(\forall k \in \mathbb{N})(\exists x \in B)(\forall n \in \mathbb{N})[n \leq k \rightarrow \neg(p_n < x < q_n)]$$

a sentence which expresses the fact that no finite subcover of \mathcal{C}' exists. By transfer,

$$(\forall k \in {}^*\mathbb{N})(\exists x \in {}^*B)(\forall n \in {}^*\mathbb{N})[n \leq k \rightarrow \neg(p_n < x < q_n)]$$

Take an unlimited $K \in {}^*\mathbb{N}$. Then there exists $x \in {}^*B$ such that $p_n < x < q_n$ is false for all standard n . Thus x cannot be infinitely close to any element of B , as then it would be in some (p_n, q_n) , so Robinson's criterion fails. \square

Corollary. (*Heine - Borel*) *A set $B \subseteq \mathbb{R}$ is compact iff it is bounded.*

Proof. If B satisfies Robinson's Criterion, it must be closed and bounded. Conversely, if x is closed and bounded, there exists a real b such that

$$(\forall x \in B)(|x| < b)$$

Let $x \in {}^*B$. By transfer, $|x| < b$. Hence x is limited, and has a real shadow r . Then $r \sim_{\mathbb{I}} x$ and $x \in {}^*\mathbb{B}$, so $r \in B$, because B is closed. □

Corollary. *The continuous image of a compact set is compact*

Proof. Let f be continuous, and let B be a compact subset of \mathbb{R} . Then

$$(\forall y \in f(B))(\exists x \in B)(y = f(x))$$

By transfer, if $y \in {}^*f(B)$, then $y = f(x)$ for some $x \in {}^*B$. Since B is compact, $x \sim_{\mathbb{I}} r$ for some $r \in B$. By continuity, $y = f(x) \sim_{\mathbb{I}} f(r) \in f(B)$. □

Corollary. *If f is continuous on compact $B \subseteq \mathbb{R}$, then f is uniformly continuous on \mathbb{R} .*

Proof. Identical to the earlier non-standard proof that a continuous function on a closed interval is uniformly continuous. □

6.6. Tychonoff's Theorem. (The product of compact topological spaces is compact)

We can generalize Robinson's criterion to arbitrary topological spaces. In that setting, we replace the argument using the density of the rationals with the following argument: Assume A is not compact, so that there exists a collection $\mathcal{U} = \{\mathcal{O}_\alpha : \alpha \in I\}$ of open sets with no finite subcover. We can analogously define the hyperreal extension ${}^*\mathcal{U}$ of a set in a topological space, with the halo of a point being

$$\text{halo}(x) = \bigcap_{U \in \mathcal{T}: x \in U} {}^*U$$

Let \mathcal{B} be a hyperfinite collection in ${}^*\mathcal{U}$ with ${}^*\mathcal{O}_\alpha \in \mathcal{B}$ for each $\alpha \in I$. Then there is a $y \in {}^*A$ such that $y \notin U$ for each $U \in \mathcal{B}$. This point y is not in the halo of any point in A , since for each $x \in A$, there exists an α with $x \in \mathcal{O}_\alpha$, but $y \notin {}^*\mathcal{O}_\alpha$.

Theorem. (*Tychonoff's Theorem*) *The product of compact spaces is compact*

Proof. Let

$$X = \prod_{\alpha \in I} X_\alpha$$

be the product of compact spaces. Let $g \in {}^*X$. Then for each standard $\alpha \in I$, there is an $x_\alpha \in X_\alpha$ with $g(\alpha) \sim x_\alpha$ (the x_α are unique if the spaces X_α are Hausdorff). The element $f \in X$ with $f(\alpha) = x_\alpha$ for each $\alpha \in I$ is in X and g is in the halo of f . \square

fn
Thanks for reading!