The Hyperreal Numbers and Applications to Elementary Calculus

Doug Babcock

25 September 2008

Abstract

The real numbers have no infinitely large or infinitely small elements. In some ways this makes them not well suited for calculus—one cannot work directly from the intuition for limits or integrals, because it's impossible to talk about things being infinitely close, or adding up the areas of rectangles that are infinitely thin. The hyperreal numbers are an extension of the real numbers in which infinitesimal elements and infinitely large elements are present; thus, in the hyperreal numbers, it is possible to approach calculus in a way which agrees with the intuition very literally. In this talk I will discuss the construction of the hyperreal numbers; some of their properties; and how to use them to prove some theorems from elementary calculus.

1 Historical Introduction

In the late 1600s, when Gottfried Wilhelm Leibniz originally developed his ideas of calculus, he thought of things in terms of infinitesimals. This is reflected in his notation: "$\int_a^b f(x)\,dx$" intuitively means the sum of the areas of rectangles, one for each real number between a and b, where the height is $f(x)$ and the width is an infinitesimal change in x. Something similar can be said about his notation $\frac{dy}{dx}$.

For more than one hundred years, this was the prevailing intuition about calculus; for example, Leonhard Euler did most of his work by thinking about infinitesimals as actual numbers. However, in the 1800s, there was a move towards making calculus rigorous. Mathematicians such as Augustin Cauchy and Karl Weierstraß looked into the idea of using infinitesimals to define limits, derivatives, and integrals, but they were unable to come up with a rigorous way of defining them; so the ϵ-δ definitions became widely used instead. In many ways, ϵ-δ definitions do not agree as directly with the intuition; they skirt the idea of infinity and infinitesimals by using ideas like "for all sufficiently large values of n." However, due to the lack of a rigorous development of a number system with infinitesimals and infinitely large numbers, references to these ideas largely disappeared from rigorous mathematics.

In the early 1900s, the work of many great logicians, such as Leopold Löwenheim, Kurt Gödel, and Thoralf Skolem, began to suggest that the real numbers were not the only structure with which it was possible to do calculus. In 1934, Skolem proved that there are structures larger than $\mathbb{N}$ which have exactly the same first-order properties that $\mathbb{N}$ does; and it is a direct corollary that the same is true about $\mathbb{R}$.

In 1961, Abraham Robinson gave the first rigorous definition of an extension of $\mathbb{R}$ which includes infinitesimals. His ideas were quickly refined by other mathematicians, such as W.A.J. Luxemburg [2], to make them more accessible. These new ideas led H. Jerome Keisler to publish a calculus textbook [1] using the new theory of infinitesimals in 1976. In this talk, I will develop Luxemburg's construction of the hyperreal numbers, and describe some of the ways in which this construction can be applied to calculus.

2 Filters

Let S be an infinite set.

Definition. A collection of subsets $\mathcal{F}$ of S is called a *filter* if it satisfies the following:
(i) $S \in \mathcal{F}$.
(ii) $\emptyset \notin \mathcal{F}$.
(iii) $A \in \mathcal{F}, A \subset B \implies B \in \mathcal{F}$.
(iv) $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$.

The intuition behind a filter is that the filter contains very large subsets of S—subsets so large that their complements are negligible.

Example. Let X be a measure space. Then the collection of subsets of X which have full measure (i.e. with complement of measure zero) form a filter on X.

Example (Principal Filter). Let $A \subset S$ be nonempty. Then $\{B \subset S \mid B \supset A\}$ is a filter on S, called the *principal filter* generated by A.

Example (Fréchet Filter). The collection $\{A \subset S \mid S \backslash A \text{ is finite}\}$ is a filter on S, called the *Fréchet filter.*

Definition. A filter $\mathcal{F}$ is called *free* if $\bigcap_{A \in \mathcal{F}} A = \emptyset$.

Proposition. The Fréchet filter is free.

Proof. Let $\mathcal{F}$ denote the Fréchet filter. For each $s \in S$, let $A_s = S \backslash \{s\}$. Then $A_s \in \mathcal{F}$ for each s, but clearly $\bigcap_{s \in S} A_s = \emptyset$. Hence $\bigcap_{B \in \mathcal{F}} B = \emptyset$ also. $\square$

Definition. A filter $\mathcal{F}$ is called an *ultrafilter* if it is maximal (that is, $\mathcal{F}' \supset \mathcal{F}$ implies $\mathcal{F}' = \mathcal{F}$).

Proposition. If $\mathcal{U}$ is an ultrafilter, then if $A \cup B \in \mathcal{U}$, then either $A \in \mathcal{U}$ or $B \in \mathcal{U}$.

Proof. Seeking a contradiction, suppose that $A \cup B \in \mathcal{U}$, but $A \notin \mathcal{U}$ and $B \notin \mathcal{U}$. Let $\mathcal{F} = \{T \subset S \mid A \cup T \in \mathcal{U}\}$. This forms a filter on S, and it is clear that $\mathcal{F} \supset \mathcal{U}$. Since $B \in \mathcal{F}$, then $\mathcal{F} \supsetneq \mathcal{U}$. This is a contradiction. □

Corollary. If $A_1 \cup \cdots \cup A_n \in \mathcal{U}$, then $A_i \in \mathcal{U}$ for some $i \in \{1, \ldots, n\}$.

Proposition. A filter $\mathcal{F}$ is an ultrafilter if and only if for all $A \subset S$, we have either $A \in \mathcal{F}$ or $S \backslash A \in \mathcal{F}$.

Proof. Suppose that $\mathcal{F}$ is an ultrafilter. Seeking a contradiction, suppose $A \notin \mathcal{F}$ and $S \backslash A \notin \mathcal{F}$. Then by the previous proposition, $A \cup (S \backslash A) \notin \mathcal{F}$, so $S \notin \mathcal{F}$, contradicting the definition of a filter.

Suppose that for every A, $A \in \mathcal{F}$ and $S \backslash A \in \mathcal{F}$. Seeking a contradiction, suppose $\mathcal{F}' \supsetneq F$ is a filter. Then $\mathcal{F}'$ contains both some set and its complement; hence it contains their intersection $\emptyset$, again contradicting the definition of a filter. □

Proposition. Every filter $\mathcal{F}$ may be extended to an ultrafilter.

Proof. Zorn's lemma. □

Proposition. Let $\mathcal{F}$ be the Fréchet filter. If we extend $\mathcal{F}$ to an ultrafilter $\mathcal{U}$, then $\mathcal{U}$ is also free.

Proof. $\bigcap_{A \in \mathcal{U}} A \subset \bigcap_{A \in \mathcal{F}} A = \emptyset$. □

3 $\mathbb{R}^*$ and its Properties

From now on, fix an ultrafilter $\mathcal{U}$ on the positive integers $\mathbb{N}$, where $\mathcal{U}$ extends the Fréchet filter. Let $\mathbb{R}^\omega$ denote the set of all sequences of real numbers. Elements of $\mathbb{R}^\omega$ will be denoted by normal lowercase letters (a, b, c, etc.); the nth element in the sequence a will be denoted by a_n.

Definition. Let $a, b \in \mathbb{R}^\omega$. We say that a and b are *equivalent modulo* $\mathcal{U}$, written $a =_{\mathcal{U}} b$, if $\{n \in \mathbb{N} \mid a_n = b_n\} \in \mathcal{U}$. Note that if a and b agree on any set in $\mathcal{U}$, then $\{n \in \mathbb{N} \mid a_n = b_n\} \in \mathcal{U}$, because $\mathcal{U}$ is closed under supersets of its elements.

Example. $(3, 1, 4, 1, 1, 1, 1, \ldots)$ and $(1, 1, 1, 1, 1, 1, 1, \ldots)$ are equivalent modulo $\mathcal{U}$, because the complement of $\{1, 2, 3\}$ is in $\mathcal{U}$ since $\mathcal{U}$ extends the Fréchet filter.

Example. $(1, 2, 1, 2, 1, 2, 1, 2, \ldots)$ is either equivalent to $(1, 1, 1, \ldots)$ or $(2, 2, 2, \ldots)$, depending on whether the set of odd positive integers or the set of even positive integers is in $\mathcal{U}$.

Definition. $\mathbb{R}^* = \mathbb{R}^\omega / =_\mathcal{U}$. Elements of $\mathbb{R}^*$ will be denoted by $\overline{a}$, meaning the equivalence class containing some $a \in \mathbb{R}^\omega$. (If we take the continuum hypothesis as an axiom, then the isomorphism class of the structure does not depend on $\mathcal{U}$.)

Definition of + in $\mathbb{R}^*$. $\overline{a} + \overline{b} = \overline{a+b}$. (Note that this makes $\overline{(0,0,0,\ldots)}$ the additive identity on $\mathbb{R}^*$.)

Proof that this is well-defined. Suppose $a, a' \in \overline{a}$ and $b, b' \in \overline{b}$. Since $a, a' \in \overline{a}$, then $\{n \in \mathbb{N} \mid a_n = a'_n\} \in \mathcal{U}$. Since $b, b' \in \overline{b}$, then $\{n \in \mathbb{N} \mid b_n = b'_n\} \in \mathcal{U}$. Then the intersection $\{n \in \mathbb{N} \mid a_n = a'_n, b_n = b'_n\} \in \mathcal{U}$. But this set is a subset of $\{n \in \mathbb{N} \mid a_n + b_n = a'_n + b'_n\}$, so the latter is in $\mathcal{U}$. Thus $\overline{a' + b'} = \overline{a + b}$. □

Definition of · in $\mathbb{R}^*$. $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$. The proof that · is well-defined is the same as the proof that + is well-defined. (Note that this makes $\overline{(1,1,1,\ldots)}$ the multiplicative identity on $\mathbb{R}^*$.)

Proposition. All nonzero elements of $\mathbb{R}^*$ have multiplicative inverses.

Proof. Let $\overline{a} \in \mathbb{R}^*$ be nonzero. Define b as follows:

$$b_n = \begin{cases} 1/a_n, & \text{if } a_n \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Notice that $\{n \in \mathbb{N} \mid a_n \neq 0\} \in \mathcal{U}$, because $\overline{a}$ was assumed to be nonzero. So $a_n b_n = 1$ on this same set in $\mathcal{U}$. Hence $\overline{a}\overline{b} = \overline{(1,1,1,1,\ldots)}$. □

Definition of < in $\mathbb{R}^*$. We say $\overline{a} < \overline{b}$ if $\{n \in \mathbb{N} \mid a_n < b_n\} \in \mathcal{U}$. The proof that < is well-defined is the same as the proof that + is well-defined.

Theorem. $\mathbb{R}^*$ is a totally ordered field.

Proof. Long, but straightforward. All axioms based on +, ·, and < which hold in $\mathbb{R}$ also hold in $\mathbb{R}^*$, because they're true in every element of every sequence in the equivalence class. The only questionable part, the existence of multiplicative inverses, was justified above. □

Observation. In $\mathbb{R}^*$, the equivalence classes of constant sequences form a copy of $\mathbb{R}$. Thus $\mathbb{R}^*$ is a field extension of $\mathbb{R}$. Henceforth, equivalence classes of constant sequences will be denoted simply by a number. (For example, $\overline{(0,0,0,\ldots)}$ will be denoted by 0.)

Observation. The set of elements equivalent to 0 modulo $\mathcal{U}$ is an ideal in $\mathbb{R}^\omega$ (and it is a maximal ideal, since $\mathbb{R}^*$ is a field).

Theorem. Let φ be a well-formed formula in first-order logic (of finite length) using $\forall$, $\exists$, $\Longrightarrow$, $\neg$, 0, 1, +, ·, <, =, and variables. (In first-order logic, variables are only allowed to stand for elements of the structure $\mathbb{R}$ or $\mathbb{R}^*$.) Then φ is true in $\mathbb{R}$ if and only if it is true in $\mathbb{R}^*$ (where the symbols 0 and 1 are interpreted as the additive and multiplicative identities in $\mathbb{R}^*$).

Proof. The proof is omitted, because it is too technical and requires some knowledge of first-order logic. However, this theorem should at least seem feasible, because $\mathbb{R}^*$ was defined simply by using the definitions of $+$, $\cdot$, $<$, $=$, 0, and 1 componentwise up to changes on a set not in the ultrafilter. For a full proof of this result, see Chang & Keisler, section 4.1. $\square$

Theorem. $\mathbb{R}^*$ is non-Archimedean; that is, there exists an element ω of $\mathbb{R}^*$ satisfying all of $\omega > 1$, $\omega > 1+1$, $\omega > 1+1+1$, etc.

Proof. Take $\omega = \overline{(1, 2, 3, 4, \ldots)}$. For any $k \in \mathbb{N}$, $\omega_n \leq k$ holds true in only finitely many places, so $\omega_n > k$ on a set in $\mathcal{U}$, since $\mathcal{U}$ contains the Fréchet filter. Thus, $\omega > k$. $\square$

Definition. An element $\overline{a}$ of $\mathbb{R}^*$ is *infinitely large* (or *infinite*) if $\overline{a} > k$ for all $k \in \mathbb{N}$. An element $\overline{a}$ of $\mathbb{R}^*$ is *infinitesimal* if $|\overline{a}| < 1/k$ for all $k \in \mathbb{N}$.

Definition. Let M_0 denote the finite elements of $\mathbb{R}^*$. Let M_1 denote the infinitesimal elements of $\mathbb{R}^*$.

Proposition. M_1 is a maximal ideal in M_0, and M_0/M_1 is isomorphic to $\mathbb{R}$.

Proof. It is straightforward to verify that M_1 is an ideal. M_1 is maximal because all elements in its complement (relative to M_0) are units. The quotient is a field, and it should be clear that is is isomorphic to $\mathbb{R}$ since we have just eliminated the "new" elements we added. (A rigorous proof is not given here.) $\square$

Definition. Let $\text{st} : M_0 \to M_0/M_1 = \mathbb{R}$ be the natural homomorphism. (This is called taking the *standard part* of a finite element of $\mathbb{R}^*$; intuitively, taking the standard part of an element of M_0 gives the real number which that element is "infinitely close to.")

Proposition (without proof). The following are true about st:
(i) $\text{st}(x + y) = \text{st}(x) + \text{st}(y)$.
(ii) $\text{st}(xy) = \text{st}(x)\text{st}(y)$.
(iii) $x \leq y \implies \text{st}(x) \leq \text{st}(y)$.
(iv) $\text{st}(x) = 0 \Leftrightarrow x$ is infinitesimal.
(v) $x \in \mathbb{R} \implies \text{st}(x) = x$.

4 Applications to Elementary Calculus

Abuse of notation. Since it has been shown that the basic operations in $\mathbb{R}^*$ are well-defined, from this point forth the elements of $\mathbb{R}^*$ will be denoted simply by lowercase letters, and no distinction will be made between the equivalence classes and the representatives thereof. It should be understood that elements of $\mathbb{R}^*$ are equivalence classes, but that distinguishing between equivalence classes and their representatives is cumbersome and unnecessary.

As you read the following results, notice how they are made easier by replacing the idea of estimation dependent on a parameter with simply taking the standard part.

Definition. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then the *nonstandard extension* of f, denoted f^*, is given by

$$f^*(a) = (f(a_1), f(a_2), f(a_3), \ldots)$$

The proof that this is well-defined is the same as the proof that + is well-defined above.

Definition. A function $f : \mathbb{R} \to \mathbb{R}$ is *continuous at x* if for every infinitesimal Δx, $f^*(x + \Delta x) - f^*(x)$ is infinitesimal.

This definition is equivalent to the usual definition of continuity. In non-standard analysis in general, the idea of a number being vanishingly small (i.e. smaller than ϵ for some appropriate choices) is replaced by the idea of being infinitesimal. This usually eliminates at least one quantifier from definitions, making them easier to grasp.

Definition. A function $f : \mathbb{R} \to \mathbb{R}$ is *differentiable at x* if there exists $D \in \mathbb{R}$ such that, for every nonzero infinitesimal Δx, the following is true:

$$\text{st}\left(\frac{f^*(x+\epsilon) - f^*(x)}{\Delta x}\right) = D.$$

Again, this is equivalent to the usual definition of the derivative. (This will also apply to all future definitions of familiar concepts.)

Proof of the Product Rule. Suppose $f = uv$, where u and v are differentiable at x. Let Δx be a nonzero infinitesimal. Define $\Delta y = f^*(x+\Delta x) - f^*(x)$; define $\Delta u = u^*(x+\Delta x) - u^*(x)$; define $\Delta v = v^*(x + \Delta x) - v^*(x)$. Then we have

$$y + \Delta y = (u + \Delta u)(v + \Delta v)$$

Multiplying out the right gives

$$y + \Delta y = uv + v\Delta u + u\Delta v + \Delta u \Delta v$$

Now $y = uv$ by definition, so we may cancel these. Let us divide by Δx also:

$$\frac{\Delta y}{\Delta x} = v\frac{\Delta u}{\Delta x} + u\frac{\Delta v}{\Delta x} + \Delta u\frac{\Delta v}{\Delta x}.$$

Taking the standard part of both sides and applying usual properties gives

$$\text{st}\left(\frac{\Delta y}{\Delta x}\right) = v\frac{du}{dx} + u\frac{dv}{dx}.$$

This is true for every nonzero Δx, so the derivative is $v\frac{du}{dx} + u\frac{dv}{dx}$. $\square$

Definition. An *infinite integer* is an element a of $\mathbb{R}^*$ such that $\{n \in \mathbb{N} \mid a_n \text{ is an integer}\} \in \mathcal{U}$.

Definition. The union of $\mathbb{N}$ with the set of infinite positive integers will be denoted $\mathbb{N}^*$. (Like with $\mathbb{R}$ and $\mathbb{R}^*$, any statement of first-order logic is true in $\mathbb{N}$ if and only if it is true in $\mathbb{N}^*$.)

Proposition. The element $p = (2, 3, 5, 7, 11, 13, \ldots)$ is prime in $\mathbb{N}^*$.

Proof. Suppose that $a \in \mathbb{N}^*$ divides p. Then $\{n \in \mathbb{N} \mid a_n | p_n\} \in \mathfrak{U}$. This set is the union of $\{n \in \mathbb{N} \mid a_n = 1\}$ and $\{n \in \mathbb{N} \mid a_n = p_n\}$, so one of these sets is in $\mathfrak{U}$. So either $a = 1$ or $a = p$. $\square$

Definition. The *nonstandard extension of the sequence* $\{s_n\}$ is a sequence $\{s_n^*\}$ indexed by infinite integers with values in $\mathbb{R}^*$, given by

$$s_n = (s_{n_1}, s_{n_2}, s_{n_3}, ...).$$

Example. If $\{s_n\}$ is the standard sequence given by $s_n = n^2$. Then if we evaluate s_k^*, where $k = (1, 3, 5, 7, 9, \ldots)$, we get that $s_k^* = (1, 9, 25, 49, 81, \ldots)$.

Definition. The sequence $\{s_n\}$ *converges to* L if for every infinite integer N, $\mathrm{st}(s_N^*) = L$.

Proposition. If $a_n \to L$ and $b_n \to M$, then $a_n b_n \to LM$.

Proof. Let N be an infinite integer. Then $\mathrm{st}(a_N^*) = L$ and $\mathrm{st}(b_N^*) = M$; so $\mathrm{st}(a_N^* b_N^*) = LM$ because st is a homomorphism. $\square$

Definition. The series $\sum_{n=1}^{\infty} a_n$ *converges to* L if, for every infinite integer N, $\mathrm{st}(s_N^*) = L$, where $\{s_n\}$ is the sequence of partial sums. If the series does not converge, then $\sum_{n=1}^{N} a_n$ is still well-defined for any infinite integer N.

Proposition. $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$.

Proof. Let N be an infinite integer. Then we have

$$\sum_{n=1}^{N} \frac{1}{n(n+1)} = \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1}\right) = \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=1}^{N} \frac{1}{n+1} = \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=2}^{N+1} \frac{1}{n} = 1 - \frac{1}{N+1}.$$

Taking the standard part gives 1, the sum of the series. $\square$

References

[1] H. Jerome Keisler. *Elementary Calculus.* Prindle, Weber & Schmidt, 1976.

[2] W.A.J. Luxemburg. Non-standard analysis: Lectures on A. Robinson's theory of infinitesimals and infinitely large numbers. Lecture notes, 1962.