J.M. HENLE

Non-nonstandard Analysis: *Real* Infinitesimals

Mathematics has had a troubled relationship with infinitesimals, a relationship that stretches back thousands of years. On the one hand, infinitesimals make intuitive sense. They're easy to deal with algebraically. They make calculus a lot of fun. On the other hand, they seem impossible to nail down. They're hard to deal with intellectually. They can mask a fundamental lack of understanding of analysis.

There was hope, when Abraham Robinson developed nonstandard analysis [R], that intuition and rigor had at last joined hands. His work indeed gave infinitesimals a foundation as members of the set of hyperreal numbers. But it was an awkward foundation, dependent on the Axiom of Choice. Unlike standard number systems, there is no canonical set of hyperreal numbers.

There is a way, however, of constructing infinitesimals naturally. Ironically, the seeds can be found in any calculus book of sufficient age. At the turn of the century, it was typical of texts to define an infinitesimal as a "variable whose limit is zero" [C]. That is the inspiration for the present approach to calculus. Its infinitesimals are sequences tending to 0.

I call the system "non-nonstandard analysis" to draw attention to its misfit nature. Having infinitesimals, it is not "standard." Nor is it "nonstandard," however, as this term now has a well-defined meaning. In what follows, we manipulate sequences of real numbers. We treat them (mostly) as numbers. We add them, subtract them, and put them into functions. They aren't *numbers*, however. Trichotomy fails, for example.

The central construction in this article is a rediscovery. Its first discoverer probably was D. Laugwitz. More on this later.

Notation. We denote a sequence $\{a_n\}_{n\in\mathbb{N}}$ by a boldface $\mathbf{a}$. We permit a_n to be undefined for finitely many n. The key idea throughout is that of "*from some point on.*" An equation or inequality involving sequences will be interpreted as being true or false, depending on whether the associated equation or inequality involving the terms of the sequences is true or false from some point on. For example,

$$\mathbf{a} > 2$$

simply means "$a_n > 2$ from some point on," that is, that for some k, $n \geq k$ implies that $a_n > 2$. We use ordinary letters for real numbers. Note that a real number r can be viewed as the sequence $\mathbf{a}$, with $a_n = r$ for all n.

Equations and inequalities involving sequences are interpreted in the same way; that is, $\mathbf{a} + \mathbf{b} = \mathbf{c}$ means $a_n + b_n = c_n$ from some point on, and $\sin(\mathbf{a}) = \mathbf{b}$ means $\sin(a_n) = b_n$ from some point on. Note that for $f(\mathbf{a})$ to be defined, it is only necessary that $f(a_n)$ be defined from some point on. The sequence $\mathbf{b}$: 0, 0, 0, 1, 2, 3, 4, . . . , for example, does have a reciprocal, $1/\mathbf{b}$, because, as noted, finitely many terms of a sequence may be undefined.

If a statement P is true from some point on and statement Q is also true from some point on, then "P and Q" is true from some point on. This fact allows us to do algebra on sequences.

Suppose, for example, we have

$$\mathbf{a} + 4 = 3\mathbf{b}$$

and

$$\mathbf{c} < 45.$$

Then,

$$a_n + 4 = 3b_n, \text{ from some point on,}$$

and

$$c_n < 45, \text{ from some point on.}$$

Because the two are true together from some point on, we can add them to get

$$a_n + 4 + c_n < 3b_n + 45, \text{ from some point on.}$$

In other words,

$$\mathbf{a} + 4 + \mathbf{c} < 3\mathbf{b} + 45.$$

What we are doing, with this construction, is taking the equivalence classes of sequences under the relation "equality from some point on." To present it in that way, however, would entail excessive formalism. One of my purposes is to demonstrate that there is a fairly simple approach to infinitesimals, one that can reasonably be presented to ordinary calculus students.

Infinitely Small and Infinitely Close

Definition. A sequence $\mathbf{a}$ is *infinitely small* if $|\mathbf{a}| < d$ for all positive real numbers d. For infinitely small $\mathbf{a}$, we write $\mathbf{a} \approx 0$. If $|\mathbf{a}| < r$ for some real r, we say $\mathbf{a}$ is *finite* or *bounded*. A sequence $\mathbf{a}$ is *infinitesimal* if $\mathbf{a} \neq 0$ and $\mathbf{a} \approx 0$.

These definitions hide quantifiers. When we say that $\mathbf{a} \neq 0$, we are actually saying that there is a k such that $a_n \neq 0$ for all $n \geq k$. Similarly, this definition of $\mathbf{a} \approx 0$ is, in reality, the more familiar and complicated $\forall \epsilon > 0\ \exists k, n \geq k \Rightarrow |a_n| < \epsilon$.

Proposition 1. *Suppose* $\mathbf{a}$, $\mathbf{b} \approx 0$. *Then*
(1) $\mathbf{a} + \mathbf{b} \approx 0$.
(2) $\mathbf{a} - \mathbf{b} \approx 0$.
(3) *If* $\mathbf{c}$ *is finite, then* $\mathbf{ac} \approx 0$.
(4) *If* $|\mathbf{c}| < |\mathbf{a}|$, *then* $\mathbf{c} \approx 0$.

PROOF. Given any positive d, we know that $|\mathbf{a}| < d/2$ and $|\mathbf{b}| < d/2$. Then, $|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}| < d$. This proves Part (1). Part (2) is proved similarly.

For Part (3), because $|\mathbf{c}| < r$ for some real r, and $|\mathbf{a}| < d/r$ for all positive d, we have $|\mathbf{ac}| < d$.

For Part (4), observe that because $|\mathbf{c}| < |\mathbf{a}| < d$, $|\mathbf{c}| < d$. ∎

Note that reals are finite, so that any result concerning finite sequences applies to reals too. Part (3) of Proposition 1, for example, tells us that if $\mathbf{a} \approx 0$, then $r\mathbf{a} \approx 0$ for any real r.

Definition. For sequences $\mathbf{a}$ and $\mathbf{b}$, we say that $\mathbf{a}$ and $\mathbf{b}$ are *infinitely close*, or $\mathbf{a} \approx \mathbf{b}$, iff $\mathbf{a} - \mathbf{b} \approx 0$.

Proposition 2. *If* $\mathbf{a} \approx r$ *and* $\mathbf{b} \approx s$, *then*
(1) $\mathbf{a} + \mathbf{b} \approx r + s$.
(2) $\mathbf{a} - \mathbf{b} \approx r - s$.
(3) $\mathbf{ab} \approx rs$.
(4) If $\mathbf{a} \leq s$, then $r \leq s$.
(5) $\mathbf{a} \div \mathbf{b} \approx r/s$, *if* $s \neq 0$.

PROOF. Parts (1) and (2) follow easily from Proposition 1.

For Part (3), note that $(\mathbf{a} - r)s$, $(\mathbf{b} - s)r$, and $(\mathbf{a} - r)(\mathbf{b} - s)$ are all ≈ 0. Adding these gives us $\mathbf{ab} - rs \approx 0$.

For Part (4), if $r > s$, then we would have both $a_n \leq s$ from some pont on and $|a_n - r| < (r - s)/2$ from some point on, which is impossible.

In view of Part (3), we need only show $1/\mathbf{b} \approx 1/s$ to establish Part (5). From $s \approx \mathbf{b}$, we get $|s/2| \approx |\mathbf{b}/2| < |\mathbf{b}|$, so by Part (4), $|s/2| \leq |\mathbf{b}|$. Then $|1/\mathbf{b} - 1/s| = |(s - \mathbf{b})/\mathbf{b}s| \leq (1/|s/2|)(1/|s|)|s - \mathbf{b}|$. By Proposition 1, this is infinitesimal. ∎

This is all we need to get started.

Elementary Calculus

Definition 1. A function *f is continuous* at $x = r$, r real, iff

$$\mathbf{a} \approx r \text{ implies } f(\mathbf{a}) \approx f(r).$$

Equivalently, f is *continuous* at r if $\mathbf{\Delta x} \approx 0$ implies $f(r + \mathbf{\Delta x}) - f(r) \approx 0$.

We can also discuss continuity at a sequence $\mathbf{a}$, as opposed to a real r, but continuity on a set X corresponds to continuity at all real numbers in X. As in nonstandard analysis, continuity at all points (real numbers and sequences) in a set X is equivalent to *uniform* continuity. I will prove this later.

Proposition 3. *The sum, difference, product, and quotient (when the divisor is nonzero) of functions continuous at $x = r$ are continuous at $x = r$.*

This follows easily from Proposition 2.

Definition 2. For a function f and a real r, we say $f'(r) = d$ iff for all infinitesimal $\mathbf{\Delta x}$,

$$\frac{f(r + \mathbf{\Delta x}) - f(r)}{\mathbf{\Delta x}} \approx d.$$

Here is what the computation looks like of the standard first example: $f(x) = x^2$.

$$\begin{aligned}\frac{(x + \mathbf{\Delta x})^2 - x^2}{\mathbf{\Delta x}} &= \frac{x^2 - 2x\mathbf{\Delta x} + \mathbf{\Delta x}^2 - x^2}{\mathbf{\Delta x}}\\ &= \frac{2x\mathbf{\Delta x} + \mathbf{\Delta x}^2}{\mathbf{\Delta x}}\\ &= 2x + \mathbf{\Delta x}\\ &\approx 2x.\end{aligned}$$

Proposition 4. *If f is differentiable at r, it is continuous at r.*

Proof. Just multiply

$$\frac{f(r + \mathbf{\Delta x}) - f(r)}{\mathbf{\Delta x}} \approx d \quad \text{and} \quad \mathbf{\Delta x} \approx 0$$

to get

$$f(r + \mathbf{\Delta x}) - f(r) \approx 0. \qquad \blacksquare$$

The proofs of the differentiation rules are simple. The product rule is typical:

Proposition 5. *If the function h is the product of differentiable functions f and g, then h is differentiable and $h' = f'g + g'f$.*

Proof. Writing $\mathbf{\Delta f}$ for $f(x + \mathbf{\Delta x}) - f(x)$ and similarly for g and h, we have $f(x + \mathbf{\Delta x}) = f(x) + \mathbf{\Delta f}$, so

$$\begin{aligned}\mathbf{\Delta h} &= h(x + \mathbf{\Delta x}) - h(x)\\ &= f(x + \mathbf{\Delta x})g(x + \mathbf{\Delta x}) - f(x)g(x)\\ &= (f(x) + \mathbf{\Delta f})(g(x) + \mathbf{\Delta g}) - f(x)g(x).\end{aligned}$$

Thus,

$$\begin{aligned}\frac{\mathbf{\Delta h}}{\mathbf{\Delta x}} &= \frac{(f(x) + \mathbf{\Delta f})(g(x) + \mathbf{\Delta g}) - f(x)g(x)}{\mathbf{\Delta x}}\\ &= \frac{f(x)\mathbf{\Delta g} + g(x)\mathbf{\Delta f} + \mathbf{\Delta f}\mathbf{\Delta g}}{\mathbf{\Delta x}}\\ &= f(x)\frac{\mathbf{\Delta g}}{\mathbf{\Delta x}} + g(x)\frac{\mathbf{\Delta f}}{\mathbf{\Delta x}} + \mathbf{\Delta f}\frac{\mathbf{\Delta g}}{\mathbf{\Delta x}}\\ &\approx f(x) \cdot g'(x) + g(x) \cdot f'(x) + 0 \cdot g'(x)\\ &= f(x) \cdot g'(x) + g(x) \cdot f'(x). \qquad \blacksquare\end{aligned}$$

The proof of the Chain Rule, omitted or banished to the appendices in virtually all texts today, is easy, but we need to discuss subsequences.

Definition 3. $\mathbf{a}$ is a *subsequence* of $\mathbf{b}$ (written $\mathbf{a} \subset \mathbf{b}$) if every term of $\mathbf{a}$ is a term of $\mathbf{b}$; more precisely, if there is an increasing function, $k : \mathbb{N} \Rightarrow \mathbb{N}$ such that $a_n = b_{k(n)}$.

Proposition 6. *Let $\mathbf{a} \subset \mathbf{b}$ be given.*

(1) *If* $\mathbf{b}$ *is infinitesimal, then so is* $\mathbf{a}$.
(2) *If* $\mathbf{b} \approx r$, *then* $\mathbf{a} \approx r$.
(3) *If* $\mathbf{b}$ *satisfies a given equation or inequality, then so does* $\mathbf{a}$.

The proof of this is routine.

Proposition 7. (The Chain Rule). *If $y = f(x)$ and $z = g(y)$, then $dz/dx = (dz/dy) \cdot (dy/dx)$.*

Proof. For $\mathbf{\Delta x}$ infinitesimal, $dy/dx \approx \mathbf{\Delta y}/\mathbf{\Delta x}$, where $\mathbf{\Delta y} = f(x + \mathbf{\Delta x}) - f(x)$. By Proposition 4, $\mathbf{\Delta y} \approx 0$, so seemingly $\mathbf{\Delta z}/\mathbf{\Delta y} \approx dz/dy$, where $\mathbf{\Delta z} = g(y + \mathbf{\Delta y}) - g(y)$. But there is another way to write $\mathbf{\Delta z}$, since $g(y + \mathbf{\Delta y}) - g(y) = g(f(x) + \mathbf{\Delta y}) - g(f(x)) = g(f(x + \mathbf{\Delta x})) - g(f(x))$, so $dz/dx \approx \mathbf{\Delta z}/\mathbf{\Delta x}$.

Putting this together,

$$\frac{dz}{dx} \approx \frac{\mathbf{\Delta z}}{\mathbf{\Delta x}} = \frac{\mathbf{\Delta z}}{\mathbf{\Delta y}} \cdot \frac{\mathbf{\Delta y}}{\mathbf{\Delta x}} \approx \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

The only difficulty with this is that $\mathbf{\Delta y}$, while infinitely close to 0, may not be infinitesimal because it may equal 0 infinitely often. In that case, we can't claim that $dz/dy \approx \mathbf{\Delta z}/\mathbf{\Delta y}$, because the definition of derivative requires $\mathbf{\Delta y} \neq 0$.

But then, let $\mathbf{\Delta x}^* \subset \mathbf{\Delta x}$ be the subsequence such that the corresponding $\mathbf{\Delta y}^*$ is a sequence entirely composed of 0's. Then, $dy/dx \approx \mathbf{\Delta y}^*/\mathbf{\Delta x}^* = 0$. And, because $\mathbf{\Delta y}^* = 0$, the corresponding $\mathbf{\Delta z}^* = g(y + \mathbf{\Delta y}^*) - g(y)$ is also 0, so $dz/dx \approx \mathbf{\Delta z}^*/\mathbf{\Delta x}^* = 0$. Thus, once again,

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}. \qquad \blacksquare$$

For integration, we can use sequences of step functions.

Definition 4. *A function s is a* **step function** *on an interval, $[a, b]$, if it is defined on the interval and changes value only a finite number of times over the interval.*

Step functions are simply functions that are piece-wise constant. A sequence of step functions, $\mathbf{s}$, is a sequence where s_n is a step function from some point on.

Definition 5. *If s is a step function on $[a_0, a_n]$ with $s(x) = c_i$ on each subinterval, (a_i, a_{i+1}), $i = 0, \ldots, n-1$, then*

$$\int_{a_0}^{a_n} s(x)dx = \sum_{i=0}^{n-1} c_i(a_{i+1} - a_i).$$

Definition 6. *For a function f and an interval $[p, q]$,*

$$\int_p^q f dx = r$$

iff there are sequences of step functions $\mathbf{d} \le f \le \mathbf{u}$ *on* $[p, q]$ *with*

$$\int_p^q \mathbf{d} dx \approx r \approx \int_p^q \mathbf{u} dx.$$

As usual, $\mathbf{d} \le f \le \mathbf{u}$ simply means that $d_n \le f \le u_n$ on $[p, q]$ from some point on, and $\int_p^q \mathbf{d}\, dx$ is the sequence $\{\int_p^q d_n\, dx\}$.

The integral, r, is unique, for if $\mathbf{d}_i$, $\mathbf{u}_i$, and r_i satisfy the conditions above for $i = 1, 2$, then

$$r_1 \approx \int_p^q \mathbf{d}_1\, dx \le \int_p^q \mathbf{u}_2\, dx \approx r_2 \approx \int_p^q \mathbf{d}_2\, dx \le \int_p^q \mathbf{u}_1\, dx \approx r_1.$$

The basic theorems on integrals are easily proved. The following is an example:

Proposition 8. *If f is integrable over $[a, b]$ and $[b, c]$, then it is integrable over $[a, c]$ and*

$$\int_a^c f\, dx = \int_a^b f\, dx + \int_b^c f\, dx.$$

PROOF. Let $\mathbf{d}_{ab}$, $\mathbf{u}_{ab}$, $\mathbf{d}_{bc}$, $\mathbf{u}_{bc}$ be the step-functions witnessing the integrability of f over $[a, b]$ and $[b, c]$. Define $\mathbf{d}$ and $\mathbf{u}$ on $[a, c]$ by gluing together the respective lower and upper functions. Then, we certainly have

$$\mathbf{d} \le f \le \mathbf{u}$$

and we also have

$$\int_a^c \mathbf{d}\, dx = \int_a^b \mathbf{d}_{ab}\, dx + \int_b^c \mathbf{d}_{bc}\, dx \approx \int_a^b \mathbf{u}_{ab}\, dx + \int_b^c \mathbf{u}_{bc}\, dx = \int_a^c \mathbf{u}\, dx. \quad \blacksquare$$

Analysis

The outstanding power of Robinson's nonstandard analysis is evident in the nonstandard proof of theorems, such as the Intermediate Value Theorem, and the integrability of continuous functions. We can do that here too, and the proofs are startlingly similar. Our tool will be the non-nonstandard equivalent of "every finite nonstandard number is infinitely close to a real." The following is effectively the Bolzano–Weierstrass theorem.

Proposition 9. *If* $\mathbf{a}$ *is finite, then for some* $\mathbf{c} \subset \mathbf{a}$ *and some real* r, $\mathbf{c} \approx r$.

PROOF. As $\mathbf{a}$ is finite, there is some d such that $|\mathbf{a}| < d$. That means there are only a finite number of possibilities for the integer part of each a_n. One of these possibilities must occur an infinite number of times. Let k be such that for infinitely many a_n, the integer part of a_n is k. Let c_1 be the first term in $\mathbf{a}$ with integer part k.

Now of the infinitely many $\{a_n\}$ having integer part k, there are only 10 possibilities $(0, 1, 2, \ldots, 9)$ for the first digit after the decimal point. One of those possibilities must occur an infinite number of times. Let d_1 be such a digit. Let c_2 be the first term in $\mathbf{a}$ after c_1, which begins "$k.d_1$."

We continue in this way, finding d_2 such that infinitely many terms begin "$k.d_1d_2$" and choosing c_3 so that it begins "$k.d_1d_2$," and so on. When we are done, we have a subsequence $\mathbf{c} \subset \mathbf{a}$, and a real number, $r = k.d_1d_2d_3d_4\ldots$, and by construction,

$$|c_1 - r| < 1,$$
$$|c_2 - r| < 0.1$$
$$|c_3 - r| < 0.01$$
$$\vdots$$

Thus, for any $q > 0$, $|c_n - r| < q$ from some point on. That means $|\mathbf{c} - r| \approx 0$, or $\mathbf{c} \approx r$. $\blacksquare$

Proposition 10 (Intermediate Value Theorem). *If f is continuous on $[p, q]$ and $f(p) \le s \le f(q)$, then for some $r \in [p, q]$, $f(r) = s$.*

PROOF: I begin by constructing two sequences $\mathbf{a}$ and $\mathbf{b}$ in the interval $[p, q]$. For $n = 1$, let $a_1 = p$ and $b_1 = q$. For n in general, divide the interval $[p, q]$ by points

$$p = x_0 < x_1 < x_2 < \cdots < x_n = q,$$

equally spaced, a distance of $(q - p)/n$ apart. As $f(x_0) \le s \le f(x_n)$, there must be two adjacent points, x_k and x_{k+1} such that $f(x_k) \le s \le f(x_{k+1})$. Let a_n be the first point, x_k, and let b_n be the second.

We certainly have $f(\mathbf{a}) \le s \le f(\mathbf{b})$. We also have $\mathbf{a} \approx \mathbf{b}$. By Proposition 9, there is a subsequence $\mathbf{c} \subset \mathbf{a}$ and a real r such that $\mathbf{c} \approx r$. Let $\mathbf{d} \subset \mathbf{b}$ be the corresponding subsequence of $\mathbf{b}$. By Proposition 6, $f(\mathbf{c}) \le s \le f(\mathbf{d})$ and $\mathbf{c} \approx \mathbf{d}$. By continuity, $f(\mathbf{c}) \approx f(r) \approx f(\mathbf{d})$. Now putting everything together,

$$f(r) \approx f(\mathbf{c}) \le s \le f(\mathbf{d}) \approx f(r).$$

By Proposition 2, Part (4), $f(r) \le s \le f(r)$, so $f(r) = s$. $\blacksquare$

Proposition 11 (Extreme Value Theorem). *If f is continuous on $[p, q]$, then f attains a maximum on $[p, q]$.*

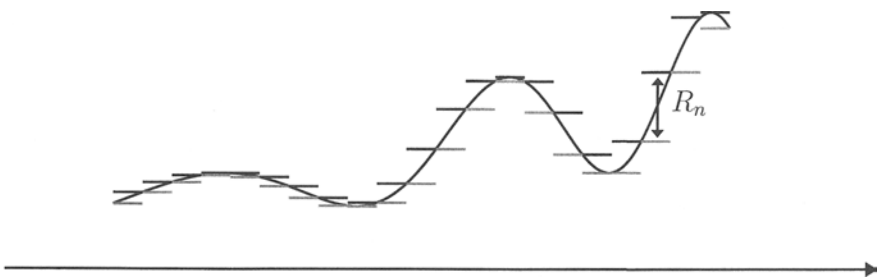
PROOF: I will define a single sequence $\mathbf{a}$ in $[p, q]$. For n, divide $[p, q]$ as in the previous proof. Let a_n be the division point x_k for which $f(x_k)$ is greatest. Choose $\mathbf{c} \subset \mathbf{a}$ and r so that $\mathbf{c} \approx r$. I claim that f reaches its maximum at $x = r$. To see this, take any s in $[p, q]$. Form sequence $\mathbf{b}$ by choosing b_n for each n to be the division point x_k nearest to s. We have that $\mathbf{b} \approx s$, as $|b_n - s| < (q - p)/n$, so $|b_n - s| < w$ from some point on for any positive w. We also have, by the construction of $\mathbf{a}$, that $f(\mathbf{b}) \le f(\mathbf{a})$.

Now, let $\mathbf{d}$ be the subsequence of $\mathbf{b}$ corresponding to $\mathbf{c}$. As before, $f(\mathbf{d}) \le f(\mathbf{c})$ and $\mathbf{d} \approx s$. So, by continuity, $f(s) \approx f(\mathbf{d}) \le f(\mathbf{c}) \approx f(r)$. It follows that $f(s) \le f(r)$. $\blacksquare$

Proposition 12. *If f is continuous on $[p, q]$, then f is integrable on $[p, q]$.*

PROOF. For any n, let l_n be the step-function formed by partitioning $[p, q]$ into n equal subintervals and setting $l_n(x)$

FIGURE 1



on each subinterval to be the minimum value of f (guaranteed to exist by Proposition 11). Similarly, define u_n as the maximum value of f. We have that $\mathbf{l} \leq f \leq \mathbf{u}$.

Because f is bounded, the sequences $\mathbf{L} = \int_p^q \mathbf{l}\, dx$ and $\mathbf{U} = \int_p^q \mathbf{u}\, dx$ are bounded, so there is a subsequence $\mathbf{l}^* \subset \mathbf{l}$ and real r such that the corresponding $\mathbf{L}^* = \int_p^q \mathbf{l}^*\, dx \approx r$. Let $\mathbf{u}^*$ be the subsequence of $\mathbf{u}$ correpsonding to $\mathbf{l}^*$, and let, for each n, Δx_n be the length of the corresponding subinterval. I claim that $\int_p^q \mathbf{u}^*\, dx \approx r$ and so $\int_p^q \mathbf{f}\, dx = r$.

Proof of claim. For any n, we can find the greatest difference, R_n, between $l_n^*(x)$ and $u_n^*(x)$ on $[p, q]$, and we have that

$$\int_p^q u_n^*\, dx - \int_p^q l_n^*\, dx \leq R_n(q - p).$$

The difference, R_n, can be represented as $|f(s_n) - f(t_n)|$, for some s_n and t_n with $|s_n - t_n| \leq \Delta x_n$.

Then,

$$\int_p^q \mathbf{u}^*\, dx - \int_p^q \mathbf{l}^*\, dx \leq \mathbf{R}(q - p) = |f(\mathbf{s}) - f(\mathbf{t})|(q - p).$$

We would like to say that as $|\mathbf{s} - \mathbf{t}| \leq \mathbf{\Delta x} \approx 0$, then, by continuity, $|f(\mathbf{s}) - f(\mathbf{t})| \approx 0$, so

$$\int_p^q \mathbf{u}^*\, dx - \int_p^q \mathbf{l}^*\, dx \approx 0 \quad \text{and} \quad \int_p^q \mathbf{l}^*\, dx \approx r.$$

But continuity requires that one of $\mathbf{s}$ and $\mathbf{t}$ be real. The property $\mathbf{s} \approx \mathbf{t} \Rightarrow f(\mathbf{s}) \approx f(\mathbf{t})$ is actually equivalent to uniform continuity. We can easily work around this, however, by finding a subsequence $\mathbf{s}^{**} \subset \mathbf{s}$ and real r, with $\mathbf{s}^{**} \approx r$ and using the corresponding subsequences $\mathbf{u}^{**}$, $\mathbf{l}^{**}$, and $\mathbf{t}^{**} \approx r \approx \mathbf{s}^{**}$ to finish the proof. ■

How Did We Do It?

How did I avoid the Axiom of Choice? I simply asked less of our sequences than one asks of numbers. They aren't totally ordered, for example. The sequences 0, 1, 0, 1, . . . and 1, 0, 1, 0, . . . are incomparable. There are also zero divisors. This doesn't cause any problems.

Why didn't I need the Transfer Principle? The Transfer Principle is a powerful schema of nonstandard analysis that says that any statement (in a particularly rich, well-defined language) that is true about the real number system is true about the hyperreal number system and vice versa. We have here a weak form of this, namely that all equations and inequalities true about reals are true about sequences. We also have conjunctions of these (but not disjunctions).

In elementary calculus and analysis, the Transfer Principle is used chiefly to prove that the nonstandard definitions are equivalent to the standard definitions. But, here, these equivalences are easy. Here's an example:

Definition 7. f *is uniformly continuous on* C *iff*

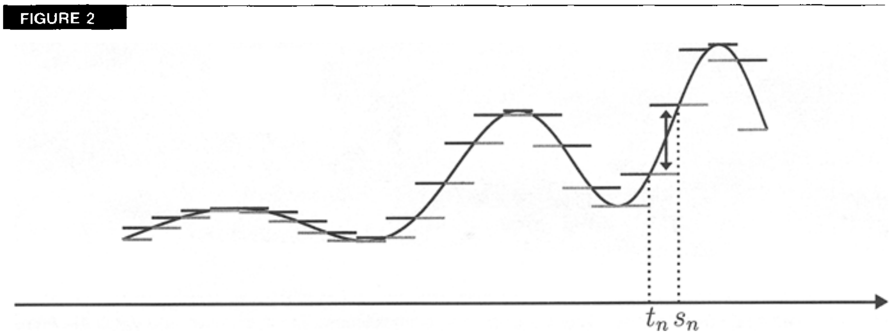
(Standard) $\forall \epsilon > 0\ \exists \delta > 0, \forall x, y \in C$
$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$

(Non-nonstandard) $\forall \mathbf{a}, \mathbf{b} \in C$, $\mathbf{a} \approx \mathbf{b} \Rightarrow f(\mathbf{a}) \approx f(\mathbf{b})$.

Proposition 13. *The standard and the non-nonstandard definitions of uniform continuity are equivalent.*

PROOF. Suppose the non-nonstandard definition holds, and suppose we are given an ϵ for which there is no suitable δ. Then, for each natural number n, choose a_n and b_n such that $|a_n - b_n| < 1/n$, but $|f(a_n) - f(b_n)| \geq \epsilon$. Then, we have $\mathbf{a} \approx \mathbf{b}$, but $f(\mathbf{a}) \neq f(\mathbf{b})$, a contradiction.

FIGURE 2



Suppose, now, that the standard definition holds, and we are given $\mathbf{a} \approx \mathbf{b}$ and d, a positive real. By the standard definition, there is a δ such that $\forall x, y\ |x - y| < \delta \Rightarrow |f(x) - f(y)| < d$. Then, because $|a_n - b_n| < \delta$ from some point on, $|f(a_n) - f(b_n)| < d$ from some point on, and so $f(\mathbf{a}) \approx f(\mathbf{b})$. ■

What happened to the Axiom of Completeness? I did use the completeness of the real line, but in a most innocuous and comprehensible form: I simply assumed that every infinite decimal corresponds to a real number.[1]

What happened to all the quantifiers? In the case of uniform continuity, for example, I went from "$\forall \epsilon\ \exists \delta\ \forall x, y \ldots$" (logicians call this a Π_3 statement) to "$\forall \mathbf{a}, \mathbf{b}, \ldots$" (a Π_1 statement). Some of the quantifiers are buried. The statement "$\mathbf{a} \approx \mathbf{b}$" is, in reality, something like "$\forall r > 0\ |a_n - b_n| < r$ from some point on," or "$\forall r > 0\ \exists k\ \forall n > k \ldots$." Essentially, I simplify definitions by coding up the most difficult part.

One can also ask, **What is next?** There are theorems of analysis where nonstandard proofs are awkward or do not exist. In the former category is the theorem that the uniform limit of continuous functions is continuous. This is proved in [HK] by taking a nonstandard model of a nonstandard model. It can be done in the present system by taking sequences of sequences. Indeed, by closing under the operation of "taking sequences," a great deal more analysis can be handled. In [He], this is pursued to prove the Baire Category Theorem, for which no simple nonstandard proof exists, and to develop measure theory.

Pedagogy

The non-nonconstructive infinitesimals presented here could improve the teaching of calculus. In both "standard" and "reform" calculus courses, rigor has been almost entirely omitted. Consequently, students are not asked to prove theorems until they have a fairly strong intuition for the subject and have met infinite sequences. This background makes non-nonstandard analysis very attractive for, say, an Advanced Calculus course, or even Calculus III.

For those wishing to remain as "standard" as possible, one can still use sequences. The chief adjustment is to replace $\mathbf{a} \approx 0$ with $\mathbf{a} \to 0$. The definitions and proofs in this article are easily converted. This approach is being used now in [CH].

Some History

The Greeks explored the ideas of infinity and infinitesimals. On the whole, they rejected them. To Aristotle, there was no absolute infinity. Completed infinite sets such as the natural numbers $\{1, 2, 3, \ldots\}$, did not exist. There could only be potential infinities; for example, for every number, there is another number that is larger, and so on. The distinction is similar to that between an infinitesimal number and the power to find ever smaller numbers. In a weak sense, it is the difference between infinitesimals and limits. Archimedes used infinitesimals for intuition [A1], then verified his results by proving them with (what we would call today) limits [A2].

Calculus, as formulated in the seventeenth century, was first expressed in terms of infinitesimals. As infinitesimals themselves were not well understood, there were criticisms and misunderstandings. On the whole, however, the doubts were overshadowed by the outstanding success of the theory as developed in the eighteenth century.

In the nineteenth century, Cauchy, Weierstrass, and others made infinitesimals unnecessary. Absolute infinities were replaced by limits, now rigorously defined. But linguistic habits didn't change. Mathematicians and physicists continued to talk in terms of infinitesimals. Infinitesimals didn't disappear from calculus texts for over 100 years.

Infinitesimals began to reappear in the twentieth century. The idea of sequences as infinitesimals appears in a remarkable book, *The Limits of Science, Outline of Logic and Methodology of Science* by Leon Chwistek, painter, philosopher, and mathematician [Ch]. Published in 1935, the book is not well known today (it is in Polish). Chwistek's definitions are similar to those presented here, although there are differences and limitations. His work foreshadows not only non-nonstandard analysis, but nonstandard analysis.

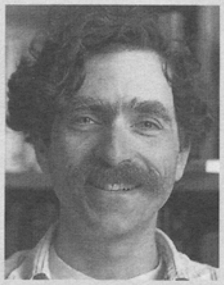
In [L1] and [L2], Laugwitz formulated the system of sequences described in this article, but with different notation. Laugwitz used his "Ω-Zahlen" to investigate distributions and operators. An earlier paper by Schmieden and Laugwitz [SL] used a more primitive system with the idea of justifying the infinitesimals of Leibniz. Laugwitz's work was not carried further, possibly because the discovery of nonstandard analysis made *real* infinitesimals unglamorous.

In 1960, Abraham Robinson constructed nonstandard models of the real number system using mathematical logic [R]. Robinson credits the papers of Laugwitz and Schmieden with some inspiration for his work. Nonstandard analysis requires a substantial investment (mathematical logic and the Axiom of Choice) but pays great dividends. Nonstandard analysis has been used to discover new theorems of analysis. It has been fruitfully applied to measure theory, Brownian motion, and economic analysis, to name just a few areas. Attempts to reform calculus instruction along infinitesimal lines, however, did not have much success [HK], [K]. Robinson's book contains an excellent history of infinitesimals.

There are other systems of standard infinitesimals, Conway's surreal numbers, for example [Co]. There are other systems for avoiding ϵ's and δ's (see [Hi] for a recent example). There may also be other rediscoveries of this system. The intended contribution of the present article is to place the structure in an algebra suitable for students of calculus.

[1] This unremarkable statement is accepted easily by students. It is not equivalent to Completeness, but is stronger, and yields the Archimedean Principle as well.

AUTHOR



J.M. HENLE
Department of Mathematics
Smith College
Northampton, MA 01063, USA
e-mail: jhenle@math.smith.edu

Jim Henle met infinitesimals as an undergraduate at Dartmouth. He studied logic at M.I.T. where he earned his doctorate in 1976. He is a professor at Smith College. His hobbies include music, gastronomy, and of course, mathematics.

Acknowledgments
I would like to thank Michael Henle, Ward Henson, Roman Kossak, Dan Velleman, and several anonymous reviewers for careful reading and helpful remarks.

REFERENCES

[A1] Archimedes, *The Method of Archimedes, recently discovered by Heiberg*, edited by Sir Thomas L. Heath, Cambridge, Cambridge University Press (1912).

[A2] Archimedes, "The quadrature of the parabola, in *The Works of Archimedes*, edited by Sir Thomas L. Heath, Cambridge: Cambridge University Press (1897).

[C] Chandler, G. H., *Elements of the Infinitesimal Calculus*, 3rd ed., New York: John Wiley & Sons (1907).

[Ch] Chwistek, L., *The Limits of Science, Outline of Logic and Methodology of Science*, Lwów-Warszawa: Ksiażnica-Atlas (1935).

[CH] Cohen, D., and Henle, J. M., *Conversational Calculus*, Reading, MA: Addison-Wesley/Longman (in press).

[Co] Conway, J. H., *On Numbers and Games*, New York: Academic Press (1976).

[He] Henle, J. M., Higher non-nonstandard analysis: Category and measure, preprint.

[HK] Henle, J. M., and Kleinberg, E. M., *Infinitesimal Calculus*, Cambridge, MA: M.I.T. Press, (1979).

[Hi] Hijab, O., *Introduction to Calculus and Classical Analysis*, New York: Springer-Verlag, (1997).

[K] Keisler, H. J., *Elementary Calculus*, An Infinitesimal Approach, 2nd ed., Boston: Prindle, Weber & Schmidt (1986).

[L1] Laugwitz, D., Anwendungen unendlich kleiner Zahlen I, *J. Reine Angew. Math.* **207** (1960), 53–60.

[L2] Laugwitz, D. Anwendungen unendlich kleiner Zahlen II, *J. Reine Angew. Math.* **208** (1961), 22–34.

[R] Robinson, A., *Nonstandard Analysis*, Amsterdam: North-Holland, (1966).

[SL] Schmieden, C., and Laugwitz, D., Eine Erweiterung der Infinitesimalrechnung, *Math. Zeitschr.* **69** (1958), 1–39.