Infinitesimal Calculus

Paul Schulze
Colgate University

Ultrafilters

Let S be a set. If $\mathcal{F} \subseteq \mathcal{P}(S)$, we say that $\mathcal{F}$ is an *ultrafilter* on S if:

- For any $A, B \in \mathcal{F}$, we have $A \cap B \in \mathcal{F}$
- If $A \in \mathcal{F}$ and $A \subseteq B \in \mathcal{P}(S)$, then $B \in \mathcal{F}$
- For any $A \subseteq S$, either $A \in \mathcal{F}$ or $A^c \in \mathcal{F}$
- $\mathcal{F}$ is a proper subset of $\mathcal{P}(S)$

We will use $\mathcal{F}$ to denote an arbitrary ultrafilter on $\mathbb{N}$ that is *non-principle*, meaning (in this context) it does not contain any finite sets. Since $\mathcal{F}$ is an ultrafilter, this means any *cofinite set*, i.e. any set whose complement is finite, is in $\mathcal{F}$.

Ultrapower of $\mathbb{R}$

Let $\mathbb{R}^{\mathbb{N}}$ denote the set of sequences of real numbers. We will denote a member $r = \langle r_1, r_2, r_3, \ldots \rangle$ of $\mathbb{R}^{\mathbb{N}}$ by $\langle r_n \rangle$. We can define operations $\oplus$ and $\odot$ on $\mathbb{R}^{\mathbb{N}}$ by $\langle r_n \rangle \oplus \langle s_n \rangle = \langle r_n + s_n \rangle$ and $\langle r_n \rangle \odot \langle s_n \rangle = \langle r_n \cdot s_n \rangle$, giving us a ring $\left(\mathbb{R}^{\mathbb{N}}, \oplus, \odot\right)$.
We define an equivalence relation $\equiv$ by $\langle r_n \rangle \equiv \langle s_n \rangle$ if $\{n \in \mathbb{N} \,|\, r_n = s_n\} \in \mathcal{F}$. Write $[[r_n = s_n]] = \{n \in \mathbb{N} \,|\, r_n = s_n\}$. Now, $\equiv$ is reflexive, as $[[r_n = r_n]] = \mathbb{N}$, and $\emptyset \notin \mathcal{F}$ implies $\mathbb{N} \in \mathcal{F}$. $\equiv$ is symmetric, since $[[r_n = s_n]] = [[s_n = r_n]]$. And $\equiv$ is transitive, since if $[[r_n = s_n]] \in \mathcal{F}$ and $[[s_n = t_n]] \in \mathcal{F}$, we have both $[[r_n = s_n]] \cap [[s_n = t_n]] \in \mathcal{F}$ and $[[r_n = s_n]] \cap [[s_n = t_n]] \subseteq [[r_n = t_n]]$
We let $[r]$ denote the equivalence class of r under $\equiv$. This can also be written as $[\langle r_n \rangle]$, or abbreviated as $[r_n]$. We then define the *hyperreals* $*\mathbb{R} = \{[r] \,|\, r \in \mathbb{R}^{\mathbb{N}}\}$. Note that $*\mathbb{R}$ in some sense contains $\mathbb{R}$: we can identify any real number a with he hyperreal $[\langle a, a, a, \ldots \rangle]$.

Hyperreal extensions

We define addition on the hyperreals $[r]+[s] = [r \oplus s]$, and we define multiplication similarly. Furthermore, we can extend any function $f : \mathbb{R} \to \mathbb{R}$ to the hyperreals by letting $*f([r_n]) = [f(r_n)]$. For instance, if $f(x) = x^2$ and $[r] = [\langle 1, 1/2, 1/3, \ldots \rangle]$, then $*f([r]) = [\langle 1, 1/4, 1/9, 1/16, \ldots \rangle]$.
Better yet, we can extend any k-ary relation R_k by saying that for any $[r^1], [r^2], [r^3], \ldots, [r^k] \in *\mathbb{R}$ (that's an upper index, not an exponent), we have $*R_k([r^1], [r^2], \ldots, [r^k])$ if $[[R_k(r^1_n, r^2_n, \ldots, r^k_n)]] \in \mathcal{F}$. In particular, $[r] < [s]$ if $[[r_n < s_n]] \in \mathcal{F}$. Doing this for 1-ary relations lets us extend any set to the hyperreals. For instance, $[r] \in *\mathbb{N}$ if $[[r_n \in \mathbb{N}]] \in \mathcal{F}$. $*\mathbb{N}$ are called the *hypernaturals*.
We must, of course, prove that all of these extensions are well-defined and do not depend on the representative we choose for $[\langle r_n \rangle]$, but this follows from the properties of $\mathcal{F}$.

Transfer

We construct a mathematical language with the following symbols:

- $\wedge$ (and), $\vee$ (or), $\rightarrow$ (if-then), and $\neg$ (not)
- $\forall$ (for all) and $\exists$ (there exists)
- a symbol c for every $c \in \mathbb{R}$
- a symbol f for every function $f : \mathbb{R} \to \mathbb{R}$
- a symbol R_k for every k-ary relation $R_k \subseteq \mathbb{R}^k$ (including $=$)

Any sentence in this language about the reals we can "reinterpret" as being about the hyperreals. For instance, say $f : \mathbb{R} \to \mathbb{R}$. If we want to say that for every real number x, there is a natural number larger than $f(x)$, we write:

$$(\forall x \in \mathbb{R})(\exists y \in \mathbb{N})(f(x) < y)$$

To "reinterpret" this about the hyperreals, we replace objects with their extensions where possible:

$$(\forall x \in *\mathbb{R})(\exists y \in *\mathbb{N})(*f(x) < y)$$

Transfer Principle: Any sentence we can write in this language is true in $\mathbb{R}$ if and only if its "reinterpretation" in $*\mathbb{R}$ is true.

Structure of $*\mathbb{R}$

We call an element $x \in *\mathbb{R}$ *infinitesimal* if $|x| < r$ for any $r \in \mathbb{R}^+$ (here treating $[\langle r, r, \ldots \rangle]$ as r). The only infinitesimal real number is 0, but consider $x = [\langle 1, 1/2, 1/3, 1/4 \ldots \rangle]$. For any $r \in \mathbb{R}^+$, by the archimedian property there is some N such that $1/N < r$, and so $[[x_n < r]] \subseteq \{n \in \mathbb{N} \,|\, N \leq n\}$. This set is confinite, and so is in $\mathcal{F}$. Hence, $x < r$. However, $[[0 < x_n]] = \mathbb{N} \in \mathcal{F}$, so $0 < x$.
We call an element $x \in *\mathbb{R}$ *unbounded* if $|x| > r$ for any $r \in \mathbb{R}^+$. For example, $[\langle 1, 2, 3, 4, \ldots \rangle]$ is unbounded. An element that is not unbounded is *bounded*. An element that is bounded but not infitesimal is *appreciable*.
If $|x - y|$ is infinitesimal, we say that x and y are *infinitely close* and write $x \simeq y$. So x is infinitesimal iff $x \simeq 0$. If x is bounded, then there is a unique $r \in \mathbb{R}$ such that $x \simeq r$ called the *standard part* of x, denoted $\mathrm{st}(x)$. Note $\simeq$ is a transitive relation, so if $\mathrm{st}(x) = \mathrm{st}(y)$ we have $x \simeq \mathrm{st}(x) = \mathrm{st}(y) \simeq y$ and so $x \simeq y$. Similarly, if $x \simeq y$ then $\mathrm{st}(x) \simeq x \simeq y \simeq \mathrm{st}(y)$, so $\mathrm{st}(x) \simeq \mathrm{st}(y)$. But $\mathrm{st}(x)$ and $\mathrm{st}(y)$ are both real, and so their difference is real. Since $\mathrm{st}(x) \simeq \mathrm{st}(y)$ implies their difference is infinitesimal, and the only real infinitesimal is 0, their difference must be 0 and so $\mathrm{st}(x) = \mathrm{st}(y)$. Hence $\mathrm{st}(x) = \mathrm{st}(y)$ iff $x \simeq y$.

Derivatives

We want to denote by $f'(x)$ an infinitesimal change in $f(x)$ over an infinitesimal change in x. Let Δx be a nonzero infinitesimal change in x. The corresponding change in $f(x)$ is $f(x + \Delta x) - f(x)$. So we define

$$f'(x) = \mathrm{st}\left(\frac{f(x + \Delta x) - f(x)}{\Delta x}\right)$$

Since Δx can be any nonzero infinitesimal, this might not be well-defined. If it is, we say that f is *differentiable* at x. Otherwise, we say that $f'(x)$ is undefined. This is equivalent to the standard definition

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

Simple Proofs with Infinitesimals

The primary motivation behind infinitesimal calculus is to allow us to prove reuslts in intuitive ways. Take for instance, the Chain Rule: standardly, the proof has an intuitive outline but runs into several techincal issues. Nonstandardly, a rigorous proof is much simpler:

Theorem (Chain Rule). *If $f, g : \mathbb{R} \to \mathbb{R}$ are differentiable, then $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$.*

Proof. Let Δx be any nonzero infinitesimal, and $\Delta g = g(x + \Delta x) - g(x)$. Since $g'(x) = \mathrm{st}(\Delta g / \Delta x)$ is defined, $\Delta g / \Delta x$ is bounded, and so Δg must be infinitesimal (as a non-infinitesimal divided by an infinitesimal is unbounded). If $\Delta g \neq 0$, then

$$\begin{aligned} (f \circ g)'(x) &\simeq \frac{f(g(x + \Delta x)) - f(g(x))}{\Delta x} \\ &= \frac{f(g(x) + \Delta g) - f(g(x))}{\Delta g} \cdot \frac{\Delta g}{\Delta x} \\ &\simeq f'(g(x)) \cdot g'(x) \end{aligned}$$

So $(f \circ g)'(x) \simeq f'(g(x)) \cdot g'(x)$. But since both of these numbers are real, they must be identical. In the case where $\Delta g = 0$, we clearly have $f(g(x) + \Delta g) - f(g(x)) = 0$ and so we get $(f \circ g)'(x) = 0$ and $\Delta g / \Delta x = 0$ so $f'(g(x)) \cdot g'(x) = 0$. $\square$

Series

Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence a_i. We then define a funciton $s : \mathbb{N} \to \mathbb{R}$ by $s(n) = \sum_{i=0}^{n} a_i$. We can then extend that function $s : *\mathbb{N} \to *\mathbb{R}$, and take an unbounded $N \in *\mathbb{N}$ to find $s(N)$, which we might denote $\sum_{i=0}^{N} a_i$. We can then define

$$\sum_{i=0}^{\infty} a_i = \mathrm{st}\left(s(N)\right)$$

for any unbounded $N \in *\mathbb{N}$. As with derviatives, this may not be well-defined, in which case we leave it undefined and say $\sum_{i=0}^{\infty} a_i$ *diverges*. When $\sum_{i=0}^{\infty} a_i$ is defined, we say it *converges*. This agrees with the standard definition.

exp

As with the standard approach, we can define

$$\exp(x) = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

Let's go ahead and define an analogous 2-place function

$$e(x, k) = \sum_{i=0}^{k} \frac{x^i}{i!}$$

We then find that $\exp(x) \simeq e(x, N)$ for any unbounded $N \in *\mathbb{N}$. Now, note that

$$\begin{aligned} e(x + \delta, k) &= 1 + x + \delta + \frac{x^2 + 2x\delta + \delta^2}{2!} + \cdots + \frac{x^k + kx^{k-1}\delta + \cdots + \delta^k}{k!} \\ &= \left(1 + x + \frac{x^2}{2!} + \cdots\right) + \delta\left(1 + \frac{2x}{2} + \frac{3x^2}{3!} + \cdots + \frac{kx^{k-1}}{k!}\right) + \delta^2(\ldots) \\ &= e(x, k) + \delta e(x, k - 1) + \delta^2(\ldots) \end{aligned}$$

Let's denote the remainder $\delta^2(\ldots)$ by R_k. With some work, it can be shown that $|R_k| < \frac{|\delta|^2}{1 - |\delta|} \cdot \frac{\exp(|x|)}{2}$. So we have

$$(\forall \delta \in \mathbb{R})(\forall k \in \mathbb{N})\left(\left|\frac{e(x + \delta, k) - e(x, k) - \delta \cdot e(x, k - 1)}{\delta}\right| < \frac{|\delta|}{1 - |\delta|} \cdot \frac{\exp(|x|)}{2}\right)$$

Now, we can transfer this over to $*\mathbb{R}$, take δ to be some nonzero real number, and in place of k write N for some unbounded hypernatural. Then

$$\left|\frac{e(x + \delta, N) - e(x, N)}{\delta} - e(x, N - 1)\right| < \frac{|\delta|}{1 - |\delta|} \cdot \frac{\exp(|x|)}{2}$$

Since $e(x, N) \simeq \exp(x)$ for real x, and $x + \delta$ is also real, we conclude that in $\mathbb{R}$,

$$(\forall \delta \in \mathbb{R})\left(\left|\frac{\exp(x + \delta) - \exp(x)}{\delta} - \exp(x)\right| < \frac{|\delta|}{1 - |\delta|} \cdot \frac{\exp(|x|)}{2}\right)$$

We then transfer this back over to $*\mathbb{R}$ and plug in an infinitesimal d for δ to get

$$\left|\frac{\exp(x + d) - \exp(x)}{d} - \exp(x)\right| < \frac{|d|}{1 - |d|} \cdot \frac{\exp(|x|)}{2}$$

The right side of this is infinitesimal: $\frac{\exp(|x|)}{2}$ is appreciable and $\frac{|d|}{1 - |d|}$ is infinitesimal (as it is an infinitesimal divided by an appreciable number). Hence, the left side of this equation is also infinitesimal, meaning

$$\frac{\exp(x + d) - \exp(x)}{d} \simeq \exp(x)$$

Since $\exp(x)$ is real, this means

$$\mathrm{st}\left(\frac{\exp(x + d) - \exp(x)}{d}\right) = \exp'(x) = \exp(x)$$

References

[1] Robert Goldblatt. *Lectures on the Hyperreals.* Vol. 188. Graduate Texts in Mathematics. New York, NY: Springer, 1998. ISBN: 978-1-4612-6841-3 978-1-4612-0615-6. DOI: 10.1007/978-1-4612-0615-6.

[2] James M. Henle and Eugene M. Kleinberg. *Infinitesimal Calculus.* Cambridge, Mass: MIT Press, 1979. 135 pp. ISBN: 978-0-262-08097-2.