Analysis on the Hyperreals, Part 2

Nic Ford

3 Limits and Continuity

3.1 Convergence of Sequences and Series

Now that we've built up the machinery of the hyperreals, we'll see how to use it to do calculus. In order to do this, we're going to need to take definitions from analysis and turn them into hyperreal-based definitions. Using the Transfer Theorem, we will then be able to get results about the real numbers from arguments involving the infinitesimal and unlimited numbers that we wouldn't otherwise have access to.

The first idea that we generalize is convergence of sequence. If we have a sequence $s_1, s_2, s_3, \ldots$ of real numbers, we can think of it as a function from the natural numbers to $\mathbb{R}$, that is, a real-valued function s whose domain is $\mathbb{N}$. (The function just takes n to the n'th term of sequence.) So we get a function $^*s : {}^*\mathbb{N} \to {}^*\mathbb{R}$ which extends s, that is, whose value on every standard natural number is just the corresponding value of s. But we also get extra terms by plugging in nonstandard (i.e., unlimited) hypernatural numbers into $^*s$. These are called *extended terms* of the sequence, and unlike the standard terms, they don't have to be standard real numbers.

Proposition. *If (s_n) is a sequence of real numbers as above, then (s_n) converges to some real number a if and only if every extended term of $^*s$ is infinitely close to a.*

Proof. First suppose the sequence converges to a. That means that for every (standard) $\epsilon > 0$, there's some natural number m_ϵ such that, if $n > m_\epsilon$, then $|s_n - a| < \epsilon$. If N is some unlimited hypernatural number, then by $n > m_\epsilon$ for every standard ϵ, so by transferring all the sentences

$$\forall n \in \mathbb{N}\,(n > m_\epsilon \to |s_n - a| < \epsilon)$$

we see that $|s_N - a| < \epsilon$ for every standard ϵ, that is, s_N is infinitely close to a.

Now suppose every extended term is infinitely close to a. Pick some standard $\epsilon > 0$. The sentence

$$\exists m_\epsilon \in {}^*\mathbb{N}\,(\forall n \in {}^*\mathbb{N}\,(n > m_\epsilon \to |{}^*s_n - a| < \epsilon))$$

is certainly true: pick any unlimited hypernatural for m_ϵ. But the standard version of this sentence is exactly what we want. □

There is a feature of this proof that's worth pointing out because it will guide our further attempts to form an analogy between the real and hyperreal worlds. The proof worked by spelling out the correspondence between "for an arbitrarily large number" in the reals and "for every unlimited number" in the hyperreals, and similarly between "as close as we want" and "infinitely close." This idea — that the things that happen infinitely far out in the hyperreals are the things that happen arbitrarily far out in the reals — is a very useful perspective to have when working with these definitions. One way to think about hyperreal analysis is that it's a way of organizing the dependences between the m's and ϵ's that show up in real analysis by "pushing them out to infinity."

Standard facts about convergence can be proved using this characterization. For example:

Proposition. *A monotone bounded sequence converges.*

Proof. Let's assume the sequence (s_n) is increasing and bounded above. (The decreasing case is similar.) Pick some extended terms s_N. Say that we have $|s_n| < m$ for each standard m. Then by Transfer, $|s_N| < m$, so s_N is limited and it has a shadow a.

We claim a is equal to the least upper bound of the set of (standard) terms of the sequence. First, s_N is larger than all the standard terms by transferring the sentence

$$\forall k \in \mathbb{N}\,(\forall l \in \mathbb{N}\,(k \leq l \rightarrow s_k \leq s_l)).$$

So for every k, since $a \simeq s_N \geq s_k$ and both a and s_k are real, we have $a \geq s_k$, so a is an upper bound. Similarly, if b is any other upper bound, then by transferring

$$\forall k \in \mathbb{N}\,(s_k \leq b)$$

we have $a \simeq s_N \leq b$, so $a \leq b$ as desired. □

Some more are given in the exercises.

We can handle series in a similar way. If we have a series $\sum_{n=0}^{\infty} a_n$, we can form the partial sums $s_k = \sum_{n=0}^{k} a_n$. These form a sequence, and so we can extend them to form *hyperfinite sums* $s_N = \sum_{n=0}^{N} a_n$ for unlimited N. Then the original series converges to some b if and only if these hyperfinite sums are always infinitely close to b.

3.2 Limits of Functions and Continuity

We can give a similar account of the limit of a function. Roughly speaking, $\lim_{x\to a} f(x) = b$ means that f (or, to be precise, $^*f$) takes points close to x to points close to b:

Proposition. *Given a function f and a point a, $\lim_{x\to a} f(x) = b$ if and only if $^*f(x) \simeq b$ whenever x is infinitely close to (and not equal to) a.*

Proof. Suppose that the limit is b. Then for every $\epsilon > 0$, there's a $\delta > 0$ such that $|f(x) - b| < \epsilon$ if $0 < |x - a| < \delta$. (These are statements about $\mathbb{R}$.) If $x \simeq a$, though, $|x - a|$ is smaller than any such δ, so by Transfer, $|^*f(x) - b|$ is less than every standard ϵ, that is, $^*f(x) \simeq b$.

For the converse, assume the condition in the proposition is true and pick some $\epsilon > 0$. By transferring the sentence

$$\exists\delta > 0\,(\forall a\,(0 < |x - a| < \delta \rightarrow |^*f(x) - b| < \epsilon))$$

(which is true by picking an infinitesimal δ) we get our result. □

Continuity can then be expressed in terms of this characterization of limits by saying that f is continuous at a if $\lim_{x\to a} f(x) = f(a)$. Alternatively, stated purely in terms of hyperreals:

Corollary. *A function f is continuous at a if, for every hyperreal $x \simeq a$, we have $^*f(x) \simeq f(a)$.*

That is, continuous functions "preserve closeness" or "take nearby points to nearby points."

This gives us intuitively appealing proofs of the Intermediate and Extreme Value Theorems. Roughly speaking, the proofs both work by cutting an interval up into infinitely many infinitesimal pieces, then using transfer to deduce the result from the analogous result about a finite partition (which is obvious). Let's see how it works in detail:

Theorem (Intermediate Value Theorem)**.** *Suppose we have a continuous function f on some closed interval $[a,b]$. Then for every real number y between $f(a)$ and $f(b)$, there's some $c \in [a,b]$ with $f(c) = y$.*

Proof. For every standard natural number n, consider the partition of $[a,b]$ into n subintervals of width $(b-a)/n$. Then since there are only finitely many intervals in this partition, there's one of them where the value of f at the one endpoint is smaller than y and the value at the other is bigger. By transferring the tedious-to-write sentence that expresses this fact, we get it for a nonstandard hypernatural N as well. So there's some subinterval $[p,q]$ of width $(b-a)/N$ with ${}^*f(p) \leq y$ and ${}^*f(q) \geq y$ (or the other way around, but let's just deal with this case). But $(b-a)/N$ is infinitesimal, so $p \simeq q$, and by continuity, ${}^*f(p) \simeq {}^*f(q)$. So since y is sandwiched between two infinitely close numbers, we must have $y \simeq {}^*f(p) \simeq f(\operatorname{sh} p)$. Since y and $f(\operatorname{sh} p)$ are both real, they're equal. □

Theorem (Extreme Value Theorem)**.** *Every continuous function f on a closed interval $[a,b]$ is bounded and attains its maximum.*

Proof. Partition the interval into finitely many pieces as above; clearly there is a maximum value of f on the endpoints of the intervals. So again, by Transfer, the same is true for a partition into N pieces, where N is an unlimited hyperreal. (It is an exercise to write the exact sentence.) Say this maximum (among endpoints of the hyperfinite partition) occurs at p. We claim that $f(\operatorname{sh} p)$ is the maximum value of f on the interval. Pick some real $x \in [a,b]$. By Transfer, x is infinitely close to some such endpoint (in fact, within $(b-a)/N$) of one), say q. Then by continuity, $f(x) \simeq {}^*f(q) \leq {}^*f(p) \simeq f(\operatorname{sh} p)$, so since both are real, $f(x) \leq f(\operatorname{sh} p)$ as desired. □

4 Derivatives and Integrals

4.1 Derivatives and Differentials

Since we have a hyperreal-based way of defining limits, we get a definition of derivatives too: the derivative of the function f at the point x is

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

or in other words, $f'(x)$ is the derivative of f at x if

$$f'(x) = \operatorname{sh}\left(\frac{f(x+dx)-f(x)}{dx}\right)$$

for any infinitesimal dx.

Note that if f is differentiable at x (that is, $f'(x)$ is well-defined in the sense just described) then we can define $df = f'(x)\,dx$. Then the derivative is *literally* the quotient $df\,/\,dx$! Alternatively, we can define $\Delta x = dx$ and $\Delta f = f(x+\Delta x) - f(x)$; then $f'(x) = \operatorname{sh}(\Delta f/\Delta x)$. We call df the *differential* of f at x and we call Δf the *increment*. (Note that they are infinitely close to each other.) This idea lets us prove:

Proposition. *Differentiable functions are continuous.*

Proof. Since $\Delta f/\Delta x$ is limited ($f'(x)$ is its shadow) we get that $\Delta f = (\Delta f/\Delta x)\Delta x$ is infinitesimal. But Δf is just $f(x+\Delta x) - f(x)$, and Δx was just an arbitrary infinitesimal, so this proves the result. □

Using this technology, we can give simple, computational proofs of standard results about derivatives:

Proposition. *If f and g are differentiable at x, then:*

1. $(f+g)'(x) = f'(x) + g'(x)$
2. $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$
3. *If we instead insist that g is differentiable at $f(x)$, then $(f \circ g)'(x) = f'(x)g'(f(x))$.*

Proof. 1. The derivative of $f+g$ is:

$$\operatorname{sh}\left(\frac{\Delta(f+g)}{\Delta x}\right) = \operatorname{sh}\left(\frac{f(x+\Delta x) - f(x) + g(x+\Delta x) - g(x)}{\Delta x}\right) = \operatorname{sh}\left(\frac{\Delta f}{\Delta x}\right) + \operatorname{sh}\left(\frac{\Delta g}{\Delta x}\right).$$

2. The increment of fg is:

$$f(x+\Delta x)g(x+\Delta x) - f(x)g(x) = (f(x)+\Delta f)(g(x)+\Delta g) - f(x)g(x)$$

so we get that the derivative is:

$$\operatorname{sh}\left(\frac{\Delta f}{\Delta x}g(x) + f(x)\frac{\Delta g}{\Delta x} + \Delta f\frac{\Delta g}{\Delta x}\right)$$

and since $\Delta g/\Delta x$ is limited and Δf is infinitesimal, we have the result.

3. Since g is differentiable at $f(x)$, and $f(x+\Delta x)$ is infinitely close to $f(x)$, $g \circ f$ is defined at $x+\Delta x$. Note that

$$\Delta(g \circ f) = g(f(x+\Delta x)) - g(f(x)) = g(f(x)+\Delta f) - g(f(x)),$$

that is, it's also the increment of g for the infinitesimal Δf. So we get that

$$\frac{\Delta(g \circ f)}{\Delta f} \simeq g'(f(x)),$$

so

$$\frac{\Delta(g \circ f)}{\Delta x} = \frac{\Delta(g \circ f)}{\Delta f}\frac{\Delta f}{\Delta x} \simeq g'(f(x))f'(x).$$

□

Notice how the proof of the chain rule reflects the naive argument often given in calculus classes that you can "cancel the *df*'s" in the expression $(dg/df)(df/dx)$. This kind of thing happens a lot in nonstandard analysis: these notations were invented with infinitesimals in mind, so it makes sense that a formalism that is built to capture that intuition should work in the same way.

We give one more example of a proof like this, and put a few more in the exercises.

Theorem (Inverse Function Theorem)**.** *If f is continuous and strictly increasing or decreasing on some open interval (a, b) and differentiable at $x \in (a, b)$ with $f'(x) \neq 0$, then the inverse function $g = f^{-1}$ is differentiable at $y = f(x)$, and $g'(y) = 1/f'(x)$.*

Proof. The original function f is bijective, so g is certainly defined on the image of (a, b), and by the fact that f is increasing or decreasing, it's defined on some interval around y. If we knew that g is differentiable, we'd know its derivative by applying the chain rule to $g(f(x)) = x$, but we don't know that yet.

We first claim g is continuous at y. Take a nonzero infinitesimal Δy. If $g(y+\Delta y)$ and $g(y)$ aren't infinitely close, take a real number r lying between them. By the monotonicity of f, $f(r)$ would then be between $y+\Delta y$ and y, which is impossible.

Write $\Delta x = g(y+\Delta y) - g(y)$. Then, since $g(y) = x$, we have $g(y+\Delta y) = x+\Delta x$, and applying f to both sides gives that $\Delta y = f(x+\Delta x) - f(x)$. So in fact Δx is the increment of f corresponding to Δy, and Δy is the increment of g corresponding to Δx, so

$$\frac{\Delta g}{\Delta y} = \frac{1}{\Delta f/\Delta x}$$

since both sides are equal to $\Delta x/\Delta y$.

Since $\Delta f/\Delta x$ is appreciable, its reciprocal is limited, and so we get our result by taking shadows of both sides of the previous equation. □

4.2 Integrals and the Fundamental Theorem of Calculus

There is also an infinitesimal-based definition of integrals. What we'd like to do is cut up the interval into infinitely many infinitesimal-width slices and add up the areas of the rectangles based at each endpoint of the partition. There are several ways to make this idea precise. We present one of them here.

For standard real numbers a, b, ϵ, we define the *Riemann sum from a to b of width ϵ* to be

$$R_a^b(\epsilon) = \sum_{a \le k\epsilon < b} \epsilon(f(k\epsilon))$$

where k takes integer values. In other words, we partition all of $\mathbb{R}$ into intervals of length ϵ and take the Riemann sum over those intervals which overlap $[a, b)$.

Fix an infinitesimal Δx. Thinking of $R_a^b(\epsilon)$ as a function of ϵ, we can transfer it to ${}^*\mathbb{R}$ and define

$$\int_a^b f(x)\,dx = \mathrm{sh}({}^*R_a^b(\Delta x)).$$

Note that this definition seems to depend on Δx. We'll see that it actually doesn't matter what Δx is (for continuous functions, at least) once we prove the Fundamental Theorem of Calculus.

Most of the standard properties of integrals are easy consequences of corresponding properties of finite sums followed by a transfer to ${}^*R_a^b$. For example:

- We have $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$ because $a \le k\epsilon < b$ if and only if $a \le k\epsilon < c$ or $c \le k\epsilon < b$, and these can't both happen, so the (finite) sums add up.
- If $f \le g$, then $\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$ because that's clearly true for the finite sums.
- If f is the constant function c, then $\int_a^b f(x)\,dx = c(b-a)$: $R_a^b(\epsilon)$ is always at most $2c\epsilon$ off from $c(b-a)$ by considering where the a and b land relative to the ϵ partition, so by Transfer, ${}^*R_a^b(\Delta x)$ is infinitely close to $c(b-a)$.

This is enough to prove the Fundamental Theorem.

Theorem (Fundamental Theorem of Calculus)**.** *If f is continuous on an interval $[a, b]$, then $F(x) = \int_a^x f(t)\,dt$ is a differentiable function of x, and $F'(x) = f(x)$.*

Proof. For standard $\Delta x > 0$, we clearly have

$$\Delta x \min f \le F(x + \Delta x) - F(x) \le \Delta x \max f$$

by the standard properties of integrals mentioned above, and so

$$\min f - f(x) \le \frac{F(x + \Delta x) - F(x)}{\Delta x} - f(x) \le \max f - f(x).$$

where the max and min are taken on the interval $[x, x + \Delta x]$. We can get a similar statement for negative Δx. By Transfer, taking Δx infinitesimal gives us that F is differentiable with derivative f (since the left and right sides are infinitesimal by continuity). □

This gives us the independence of the integral on our choice of Δx when f is continuous. No matter what Δx is, F is an antiderivative of f, and since we also know that (with notation as above) $F(a) = 0$, this uniquely determines F.