The Nonstandard Approach

One drawback with the real analysis approach involving the epsilon-delta definition of the limit is its cumbersome notation. Proofs can quickly become cluttered with absolute values and inequalities that at times can be challenging to keep track of and connect to one another. By formalizing the idea of the infinitesimal to where basic algebraic techniques can be applied to it, a new system of analysis can be developed to study the same material that real analysis does while taking advantage of simpler and more concise notation.

Introduction

This introduction follows section 1.4 of

Keisler, H. (1976). *Foundations of Infinitesimal Calculus*, Prindle Weber & Schmidt.

Let us first consider two points on the parabola $f(x) = x^2$. One will be a fixed point at (x_0, y_0), while the other will lie some distance away at $(x_0 + \Delta x, y_0 + \Delta y)$ where Δx and Δy represent the horizontal and vertical distances between the two points.

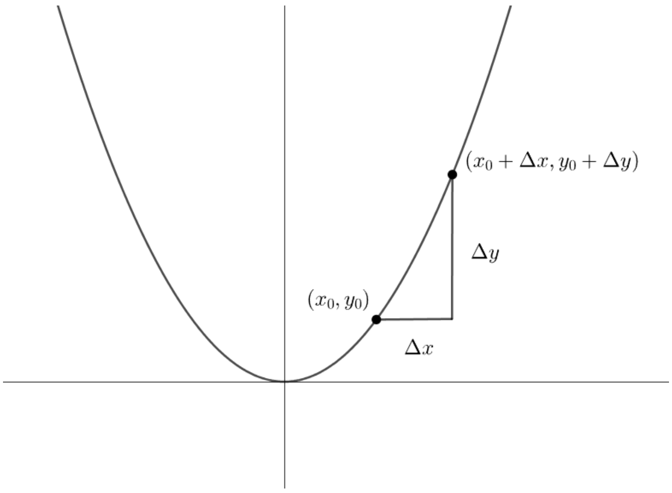

Figure 1: (x_0, y_0) and $(x_0 + \Delta x, y_0 + \Delta y)$ on $f(x) = x^2$

The average slope between any two points (x_1, y_1) and (x_2, y_2) is the ratio of the change in y to the change in x:

$$\frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}.$$

Therefore, the average slope between the two points on the parabola above is

$$\frac{(y_0 + \Delta y) - y_0}{(x_0 + \Delta x) - x_0}. \tag{1}$$

The function is $f(x) = x^2$, so any y value is determined by plugging in its corresponding input x into it. The points then become (x_0, x_0^2) and $\left(x_0 + \Delta x, (x_0 + \Delta x)^2\right)$. Substituting in these new y coordinates into (1) gives

$$\begin{aligned}
\frac{\Delta y}{\Delta x} &= \frac{(x_0 + \Delta x)^2 - x_0^2}{(x_0 + \Delta x) - x_0} \\
&= \frac{\left[x_0^2 + 2x_0\Delta x + (\Delta x)^2\right] - x_0^2}{(x_0 + \Delta x) - x_0} \\
&= \frac{2x_0\Delta x + (\Delta x)^2}{\Delta x} \\
&= 2x_0 + \Delta x. \qquad (2)
\end{aligned}$$

(2) gives us the slope in terms of a fixed point's x-coordinate and its horizontal distance Δx from some variable point. If we want to find the slope of a line that lies **tangent** to the curve – meaning that it intersects the curve at exactly one point – at (x_0, x_0^2), then we can treat Δx as a very small number so that the tangent line's slope is very close to (2). We *cannot* set Δx equal to zero, because (2) follows from a quotient where Δx is present in the denominator of a quotient. Anyhow, treating it as a minisule value leads to the intuitive result that the slope of some line tangent to the point (x_0, x_0^2) on the parabola is essentially

$$2x_0.$$

This loose usage of an infinitesimal – treating a value as infinitely small to where it can be ignored – has in fact produced decently accurate results throughout the development of calculus over the past several centuries, being used by those such as Newton, Leibniz, Euler, and others in varying forms. However, it is by no means rigorous and does not establish results with absolute certainty. Why? This is because we are left with a problem: How do we know exactly when numbers are small enough to be treated as negligible in a calculation?

This now brings us to our first formal definition:

> An **infinitesimal** is a number ε where
>
> $$-a < \varepsilon < a$$
>
> for any positive real number a.

Amongst the real numbers, 0 would be the only value that qualifies as infinitesimal. What we now do is expand the real number system by introducing **hyperreal numbers**. These include the real numbers along with nonzero infinitesimals, which can be thought of as numbers that lie infinitely close to 0. The set of all real numbers is represented by $\mathbb{R}$, while the set of all hyperreal numbers is represented by $\mathbb{R}^*$. Various symbols including $\Delta x, \Delta y, \varepsilon,$ and δ are used to symbolize infinitesimals. For instance, $x + \Delta x$ is interpretted as some quantity that lies *infinitely* close to x, and $\frac{1}{\varepsilon}$ represents an **infinite positive number**. Hyperreal numbers that are not infinitely positive or negative are **finite numbers**.

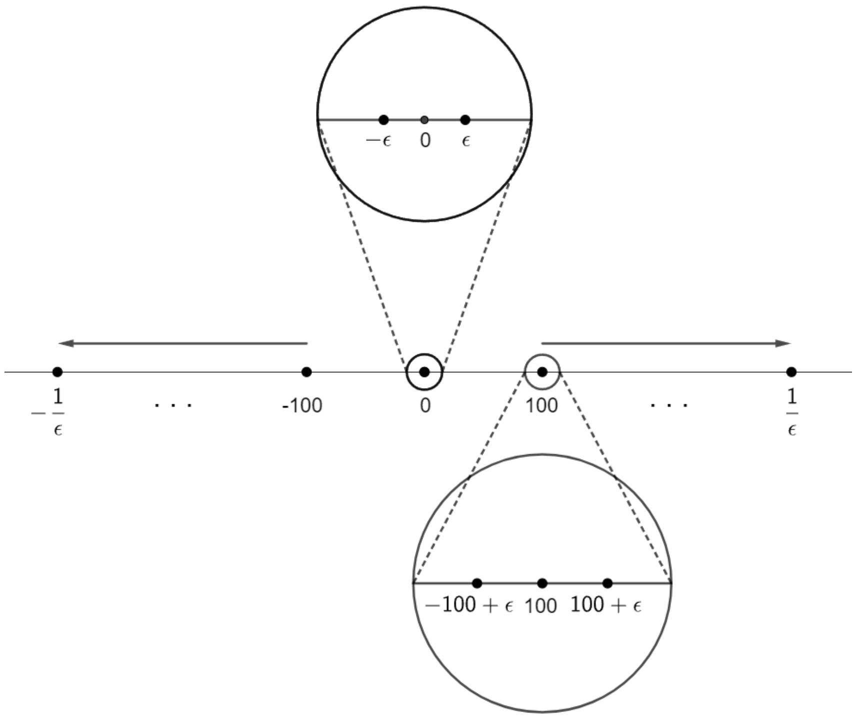

Figure 2: Hyperreal number line visualized

We can conceptualize the scale of hyperreal numbers by imagining ourselves “zooming in” infinitely close onto on a real number such as 0 or 100 on the hyperreal number line (Figure 5). This helps to clarify the nature of hyperreal numbers as merely an extension of the real numbers, so much so to where its properties and the arithmetic involved with them match closely with that of the reals.

Going back to the calculation of the instantaneous slope of $f(x) = x^2$, we now treat the value of Δx as a nonzero infinitesimal under this new lense of analysis. This means that the expression

$$2x_0 + \Delta x$$

lies infinitely close to $2x_0$, so the slope of any line tangent to (x_0, x_0^2) is concluded to be $2x_0$.

Let’s use infinitesimals to try computing the instantaneous slope of the function

$$f(x) = \sqrt{x}.$$

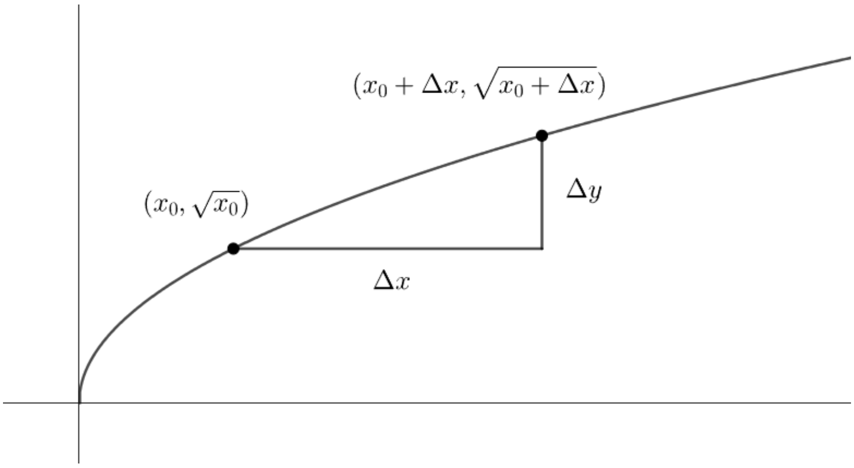

Figure 3: $f(x) = \sqrt{x}$

The average slope between a fixed point and a variable point on the square root function is

$$\begin{aligned}\frac{\Delta y}{\Delta x} &= \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \\ &= \frac{\sqrt{x_0 + \Delta x} - \sqrt{x_0}}{\Delta x}\end{aligned}$$

In order to calculate the instaneous slope, we need a way to cancel out Δx from the denominator. This can be readily accomplished by multiplying the quotient by the conjugate of the numerator in the form of 1 to induce a difference of squares:

$$\begin{aligned}\frac{\sqrt{x_0 + \Delta x} - \sqrt{x_0}}{\Delta x} \cdot \left(\frac{\sqrt{x_0 + \Delta x} + \sqrt{x_0}}{\sqrt{x_0 + \Delta x} + \sqrt{x_0}}\right) &= \frac{\left(\sqrt{x_0 + \Delta x}\right)^2 - \left(\sqrt{x_0}\right)^2}{\Delta x\left(\sqrt{x_0 + \Delta x} + \sqrt{x_0}\right)} \\ &= \frac{(x_0 + \Delta x) - x_0}{\Delta x\left(\sqrt{x_0 + \Delta x} + \sqrt{x_0}\right)} \\ &= \frac{\Delta x}{\Delta x\left(\sqrt{x_0 + \Delta x} + \sqrt{x_0}\right)} \\ &= \frac{1}{\sqrt{x_0 + \Delta x} + \sqrt{x_0}}\end{aligned}$$

Since the Δx in the denominator is infinitesimal, this means that the quantity $\sqrt{x_0 + \Delta x}$ is infinitely close to $\sqrt{x_0}$, so

$$\begin{aligned}\frac{1}{\sqrt{x_0+\Delta x}+\sqrt{x_0}} &= \frac{1}{\sqrt{x_0}+\sqrt{x_0}} \\ &= \frac{1}{2\sqrt{x_0}}.\end{aligned}$$

The skeptical reader may question how exactly this system has solved the concern regarding rigor. In particular, how do the ideas presented so far formalize the concept of deeming a quantity in a given calculation as neglible? What has happened is that we have *defined* what it means for a number to be considered negligibly small, so if a given quantity does not meet the criteria specified, then we know that it cannot be ignored. Beyond that, however, a more intimate understanding of the formalization of infinitesimals requires a working knowledge of abstract algebra. The details are obviously left out, as they lie far beyond my expertise. While the curious reader may feel discouraged or unfulfilled by this, keep in mind that the formalization of many concepts require complicated, abstract machinery, so it is commonplace for students of mathematics to take a result at face value at first and uncover its logical validity far later. As a direct analogue to the current situation, consider the analysis behind limits covered in the first half of this paper. It is likely that most calculus students who have studied limits will never encounter this rigorous treatment of them – let alone give it any consideration. Ambitious students, however, will inevitably run across it in a real analysis course further down the road, and the formalities are unveiled then. We now lie in the very same boat with infinitesimals!

Basic Applications of Infinitesimals

This section follows

> Stroyan, Keith. (2012). *A Brief Introduction to Infinitesimal Calculus*, can be found at https://homepage.math.uiowa.edu/~stroyan/InfsmlCalculus/Lecture1/Lect1.pdf

It is due time for us to see how infinitesimal arithemetic can be used to provide basic intuitive proofs of various ideas seen in calculus. Note that these are by no means formal proofs.

Theorem 4 - The Extreme Value Theorem
Consider a function $f(x)$ that is continuous on a closed interval $[a, b]$. There then exist two numbers x_{min} and x_{max} at which $f(x)$ achieves its minimum and maximum values respectively where

$$f(x_{\text{min}}) \leq f(x) \leq f(x_{\text{max}})$$

for all other x's.

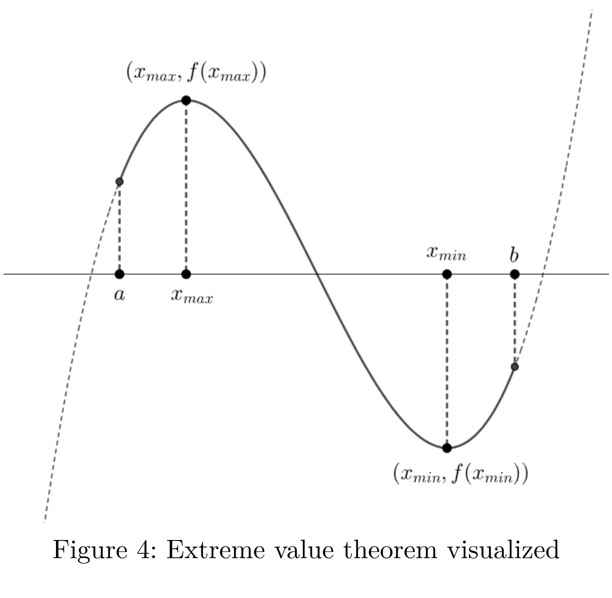

Figure 4: Extreme value theorem visualized

The extreme value theorem asserts that if a function smoothly travels over an interval of inputs, then it must take on a maximum and a minimum value at least two x's.

Continuity entails that the outputs of a function are close to one another if the inputs are close. Another way of expressing this idea is that a small change in the input should only ever produce a proportionately small change in the output. Mathematically speaking, a function $f(x)$ is continuous on a closed interval $[a, b]$ only when

$$a \leq x_1 \approx x_2 \leq b \Longrightarrow f(x_1) \approx f(x_2)$$

where the $\approx$ symbol means that two quantities are *approximately* equal.

With this property fleshed out, we can now go about providing an intuitive argument for the extreme value theorem. We start by dividing up the interval $[a, b]$ into small increments like so:

$$a < a + \frac{b-a}{H} < a + \frac{2(b-a)}{H} < \cdots < a + \frac{k(b-a)}{H} < \cdots < b,$$

where H is the number of parts the interval is partitioned into. Amongst the partition points, $f(x)$ achieves a maximum value at $x_M = a + \dfrac{k(b-a)}{H}$, so

$$f(x_m) \geq f(x_1) \text{ for any } x_1 = a + \frac{j(b-a)}{H}$$

where $j \neq k$. We know that any input x lies within the interval $[a, b]$ is within a distance of $\dfrac{1}{2H}$ from one of the partition points. To see why, consider the closed interval [0,10] divided into 5 equal subintervals ($H = 5$), so the partition points are

$$0, 0 + \frac{10-0}{5}, 0 + \frac{2(10-0)}{5}, 0 + \frac{3(10-0)}{5}, 0 + \frac{4(10-0)}{5}, 0 + \frac{5(10-0)}{5}$$