Why limits?

Math 120 Calculus I

Clark University, fall 2007

Why limits? We'll spend the next few weeks studying "limits." Naturally, the question is "why limits?" Why not just go on to derivatives?

The answer involves the character of the course. This is not just a course about how to use calculus, but a mathematics course about what calculus is.

The greatest minds of the of the 17th century, Newton and Leibniz, spent considerable time not just inventing calculus, but trying to figure out what it is. Newton and Leibniz created rules for dealing with derivatives and integrals, rules that lead to the word "calculus" for the whole subject, but neither had a very good understanding of the basis of their theory.

Leibniz rested his calculus on the concept of infinitesimals. Infinitesimals were supposed to be positive quantities less than any positive number. His theory required not just infinitesimals, but infinitely many orders of infinitesimals. He needed second-order infinitesimals smaller than any of the first-order infinitesimals, and third-order infinitesimals, and so forth. The foundations of Leibniz' infinitesimals were not logically justified until the middle of the 20th century.

Newton founded his calculus on intuitive concepts of limits. He did not define what limits were, nor did he state the properties that he expected limits to have. The foundations of Newton's limits were not logically justified until the 19th century.

What we're going to do is develop foundations of limits, the kind that Newton used only intuitively. That is going to take some time.

How do derivatives depend on limits? Just how are derivatives supposed to depend on limits? A derivative is supposed to be the rate of change of a function at an instant, what we'll call "instantaneous rate of change." On the face of it, that makes no sense at all, since an instant is a point in time, and nothing changes at a point in time. You need a time interval for anything to change.

Leibniz got around that by taking a point in time to have length, although for him it was an infinitely short length. Newton didn't do it that way. Newton recognized that an instantaneous rate of change could be found as a limit of rates of change over shorter and shorter intervals. In fact, Newton wasn't the first to use that idea since Fermat and others before Newton had already done that. Here's how they did it.

Start with a function $y = f(x)$. To help us understand, let's take x to be time, measured in some convenient time unit, and let's take $y = f(x)$ to be the distance travelled at time x, measured in some convenient distance unit. Then the derivative is what we know as velocity. The velocity doesn't have to be constant, but may change over time. It might slowly at first with a small velocity, and later quickly with a large velocity, or vice versa. We're trying to determine how to find the velocity $f'(x)$ when we know the distance $f(x)$. Finding the derivative $f'(x)$ when you know $f(x)$ is called *differentiation*.

Average rates of change. We can fairly easily compute the average rate of change, that is, the average velocity, over an interval. Suppose we take the time interval $[a, b]$ which starts at time $x = a$ and ends at time $x = b$. We can compute the distance travelled over that interval as the difference $f(b) - f(a)$. But the length of the time interval is $b - a$. That says the object travels a distance of $f(b) - f(a)$ units over a time interval of length $b - a$. Therefore, the average rate of change is

$$\frac{f(b) - f(a)}{b - a}.$$

Instantaneous rates of change. Newton recognized that the instantaneous rate of change at an instant $x = a$ can be defined as the limit of the average rates of change over an interval $[a, b]$ as the length of the interval goes to 0, that is, as b approaches a. Long after Newton, a special notation has evolved to abbreviate this expression. Namely,

$$\lim_{b \to a}$$

stands for "the limit as b approaches a." With that notation we can symbolically express the instantaneous rate of change at $x = a$ as

$$f'(a) = \lim_{b \to a} \frac{f(b) - f(a)}{b - a}.$$

Note for the limit $\lim\limits_{b \to a}$ that b is never taken to equal a. That would lead to division by 0.

But what, exactly is a limit? Newton didn't have the limit notation, but he did define derivatives in terms of limits. What Newton didn't have was a precise definition of limits. That's what we'll develop at the beginning of chapter 2. Once we have that definition, we'll be able to find various properties of limits. When we finish chapter 2, we'll define derivatives using them in chapter 3.