Introduction to limits

12 September 2017

And what are these Fluxions? The Velocities of evanescent Increments? And what are these same evanescent Increments? They are neither finite Quantities nor Quantities infinitely small, nor yet nothing. May we not call them the Ghosts of departed Quantities?

Bishop Berkeley, 1734
from *The Analyst: A Discourse Addressed to an Infidel Mathematician*

1 Introduction

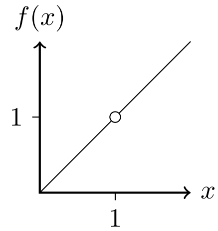
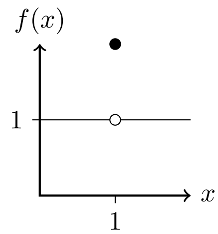
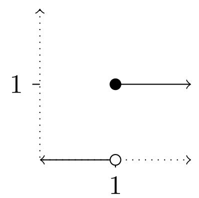
The idea of a *limit* lies at the logical foundation of calculus. Practically speaking, it is not necessary to think about limits when using calculus to solve real-world problems; indeed the full formalism of limits was not developed until after calculus had been practiced and applied successfully for hundreds of years. Nevertheless, it is an inspired idea that brings clarity to many ideas that would otherwise be fuzzy and vague. The quotation above is an excerpt from one of the an early critiques of calculus, when the subject was still done in an informal manner (complete with phrases like "infinitely small" that had no precise meaning).

Limits are a mathematical idealization. They address situations like we saw in the last lecture, where we know that the tangent line is "like" a secant line between extremely near points, yet none of these secants is precisely the tangent. I like to view them in terms of experiments: while every experiment has error, we can always shrink this error by using more precise instruments and performing the experiment more precisely. The theoretically perfect measurement, however unattainable in the real world, is what we call the *limit.* It is the sort of Platonic form of a well-done experiment.

This lecture discusses what exactly the limit concept idealizes, and presents various examples and thought experiments to show what it does and does not mean, as well as the connection to the tangent lines we discusses last time. We will return to limits in more depth as the course goes on.

2 Dot and circle notation

To highlight exactly what we mean by limit, we will sometimes need to discuss some very bizarre functions. These functions are mainly used in thought experiments, but functions like this do sometimes arise, for example, in computer science (where the subtleties of how computations are performed result in strange "edge cases"). I will describe the notation with the following three examples.

When I draw:	**I mean:**
	$f(x) = x$ for all values of x *except* $x = 1$. The value $f(1)$ is not defined.
	$f(x) = 1$ for all values of x *except* $x = 1$; $f(1) = 2$. This is often written: $f(x) = \begin{cases} 1 & \text{if } x \neq 1 \\ 2 & \text{if } x = 1 \end{cases}$.
	$f(x) = \begin{cases} 0 & \text{if } x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$ (this is called the *Heaviside function*)

So empty circles $\circ$ mean that the graph is "punctured" at the indicated point (and *only* at the indicated point) while filled circles $\bullet$ show where a value of the function has been "inserted" in place, despite its surroundings.

For a very simple example of where these sorts of graphs might be used, consider the function $f(x) = \frac{x^2-x}{x-1}$. Then $f(1)$ should be $\frac{0}{0}$, which is undefined, yet for all other values of x, this is equal to $\frac{x(x-1)}{x-1} = x$. So the graph looks like the first example above.

It is a bit confusing at first to think that $f(x) = x$ and $f(x) = \frac{x^2-x}{x-1}$ could have different graphs; after all the first function is just a "simplification" of the first. Of course this is true; the point is that simplifying a function does not give exactly the same function; it "fills in" some of the holes in the graph be eliminating the cases that look like $0/0$. This is one reason why you should be careful to simplify expressions before programming them into a computer: a computer will have an error if it attempts to divide 0 by 0. You can think of the empty circles $\circ$ as indicating inputs that would cause a computer to crash (unless the gap is "filled in" by a black dot elsewhere).

3 Limits as ideal measurements

As I said in the introduction, you can think of values of a function as outcomes of an experiment. Any real world experiment has some error. In the case of filling bottles: you can never measure *exactly* $1mL$ of water into a bottle; the best you can do is measure an amount that is equal to $1mL$ up to the accuracy of your measuring instrument.

Every time we discuss a limit in mathematics, you can think of this as meaning: we are measuring the value of a function at a point, but we are assuming that we can't actually put the *exact* input into the function that we care about. So we expect some error in the output. The *limit* is the idealization of our measurement.

To illustrate this idea, imagine that at the end of this experiment, we will measure the result using a measuring stick. The measuring stick has tick marks that are very close together (say a thousandth of an inch apart). So we can measure the result to within a thousandth of an inch. As long as we perform the experiment carefully enough (i.e. put a value into the function that is very close to the value we care about), we'll always measure the same result on the yardstick. This value is the limit, at least to accuracy of a thousandth of an inch. However, we could then go and get a new measuring stick, capable of measuring the result to within a millionth of an inch. We might notice that now our measurements are not always the same, because we are measuring more accurately. But we can respond by performing the experiment more carefully, putting much more precise inputs into the function, until we always measure the same result, to a millionth of an inch.

The process described above could go on forever. The limit is the exact, ideal value that we appear to be measuring, first to a thousandth of an inch, then to a millionth of an inch, and so on. We never measure it exactly, which is why it is an idealization.

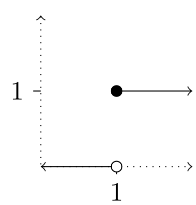
For example, suppose we are measuring the slope of a tangent line. We measure it by drawing a secant line between two very close points. If we make sure the points are close enough, and measure to some given accuracy, we'll always measure the same slope; it is exactly like we're measuring the slope of the tangent line. The more accurate we want the slope, the closer we must be sure to move the points. But we can never move the points all the way together without dividing 0 by 0. In this case, drawing a secant line is an imperfect experiment, and the slope of the tangent line is the idealized value of this experiment.

We will use the following notation and terminology for limits. The number 1 can be replaced with any other number.

- $\lim_{x \to 1} f(x)$ means the ideal measurement of $f(x)$, when we put in values of x close (but not equal) to 1.
- $\lim_{x \to 1^+} f(x)$ means the ideal measurement of $f(x)$, when we put in values of x *larger than* (but not equal to) 1. This is also called the limit limit *from the right.*
- $\lim_{x \to 1^-} f(x)$ means the ideal measurement of $f(x)$, when we put in values of x *smaller than* (but not equal to) 1. This is also called the limit *from the left.*

The first type of limit (the usual sort we will consider) is sometimes called a *two-sided limit*, while the second two are called *one-sided limits.*

Example 3.1. Consider the third example function from the previous section.



Then for this function:

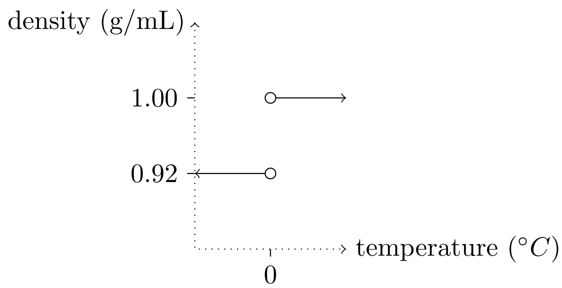
- $f(1) = 1$, by definition.
- $\lim_{x \to 1^-} f(x) = 0$, because $f(x) = 0$ for all values of x less than 1, no matter how close. **Note that the fact that $f(1) = 1$ is irrelevant to the limit.**
- $\lim_{x \to 1^+} f(x) = 1$, because $f(x) = 1$ for all values of x greater than 1. Here again, **this is unrelated to the fact that this is also the actual value** $f(1)$.
- $\lim_{x \to 1} f(x)$ **does not exist** because the two one-sided limits are not equal.

In general, the two-sided limit only exists if both one-sided limits exist, and are equal to each other.

Example 3.2. **Phases of water.** One unusual property of water is the fact that it expands when it freezes. The result of this is that the same substance (H_2O) drops somewhat in density when its temperature drops below $0°C$. Liquid H_2O has density $1.00g/mL$[1]. Solid H_2O (ice) has density $0.92g/mL$ at temperatures near $0°C$. Let $d(t)$ be the following function: it is the density of H_2O, in g/mL, when its temperature is t degrees Celcius (at standard pressure). This is an example of a function with two different two-sided limits. In fact:

- $\lim_{x\to 0^-} d(t) = 0.92$. Any experiment conducted on H_2O at temperature below (but close to) 0 will return density essentially 0.92.
- $\lim_{x\to 0^+} d(t) = 1.00$ Any experiment conducted on H_2O at temperature above (but close to 0 will return density essentially 1.00.
- $\lim_{x\to 0} d(t)$ does not exist, because if you attempt to conduct an experiment on H_2O at $0°C$, it might be solid, liquid, or some of both; the measured density in such experiments will always vary between 0.92 and 1.00, **no matter how accurately you conduct the experiment**.

In fact, this function is an example of a real-world function that is **not defined at an isolated point.** This is because $d(0)$ has no well-defined value. The reason is that if H_2O has temperature $0°C$, it could be either solid or liquid, and these two have different densities. So there is no reasonably value $d(0)$. So the graph of $d(t)$ looks something like the following (in fact, both "ends" should be slightly curved, but I have not tracked down the actual experimental data on this topic).



Example 3.3. **Absolute zero** Suppose that $f(t)$ is some function of temperature, where t is the temperature in Kelvin. Then the value $f(0)$ is not accessible by experiment, since it is impossible to actually reach absolute 0. Similarly, $f(x)$ is not accessible for any negative value of x. We *can*, however, attempt to measure $\lim_{x\to 0^+} f(t)$. To do this, we measure $f(t)$ for very small temperatures t, brought down as close as possible to $t = 0$. By doing this, and bringing the experiment apparatus better and better, we could approximate this limit to more and more decimal places. So for physical quantities depending on temperature (in Kelvin), the limit from the right makes sense, but neither the value at 0, not the limit from the left, can be accessed by physical experiment.

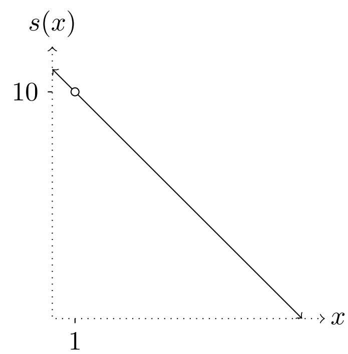
4 Examples

In this section, we discuss some examples of limits, for functions given explicitly by formulas, or from dot and circle pictures.

[1]The density of liquid water varies slightly with temperature, but below $30°C$ or so, its density is 1.00 when rounded to the nearest hundredth.

Example 4.1. (from the last lecture) What is the slope of the tangent line to the graph of $y = -x^2 + 12x - 11$ at $(1, 0)$?

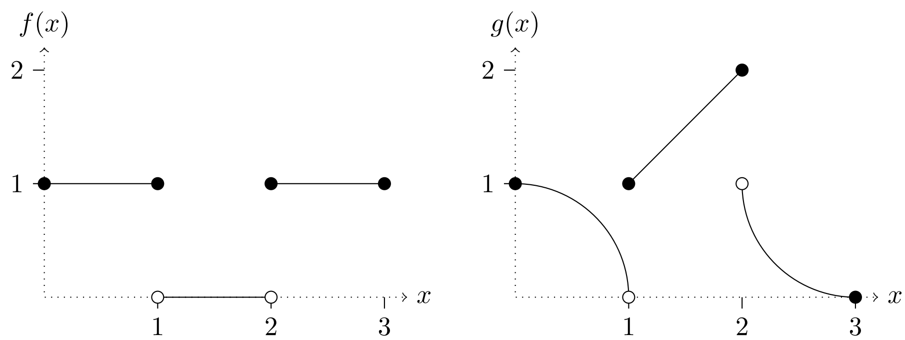
Answer. The tangent line is like an ideal secant line. As we saw in the previous lecture, the slope of the secant line from $(1, 0)$ to (x, y) (where $y = -x^2 + 12x - 11$) is given by the function $s(x) = \frac{-x^2+12x-11}{x-1}$. This expression simplifies to $11 - x$, so in fact $s(x) = 11 - x$ for all $x \neq 1$. But $s(x)$ is not defined at $x = 1$ since you cannot draw a secant from a point to itself.



Just as the secant lines approach the tangent line, the slope $s(x)$ approaches the slope of the tangent line as x approaches 1. So the slope of the tangent line is $\lim_{x\to 1} s(x)$. This limit is equal to 10, because all experiments near $x = 1$ (but not exactly 1) give slope $11 - x$, which goes to 10 as x goes to 1.

The next example consists of some thought experiments about a couple rather contrived and exotic functions. You should mainly view it at as a way to probe for intuition about what strange things are possible for rather unusual functions.

Example 4.2. Consider the following two functions, $f(x)$ and $g(x)$.



(a) For which values of c between 1 and 3 does $\lim_{x\to c} f(x)$ note exist? What about $\lim_{x\to c} g(x)$?

(b) For which values of c between 1 and 3 does $\lim_{x\to c} (f(x) + g(x))$ exist?

(c) Draw the graph of the product $f(x)g(x)$.

Solution.

(a) At $x = 1$, the limit of $f(x)$ from the left is 1, but the limit from the right is 0. So the limit $\lim_{x\to 1} f(x)$ does not exist. Similarly, $\lim_{x\to 2^-} f(x) = 0$ and $\lim_{x\to 2^+} f(x) = 1$, so $\lim_{x\to 2} f(x)$ does not exist since these are different. At all other points, the two one-sided limits exist and coincide.

Similarly, for $g(x)$, the picture shows that:

$$\begin{array}{ll} \lim_{x\to 1^-} g(x) = 0 & \lim_{x\to 2^-} g(x) = 2 \\ \lim_{x\to 1^+} g(x) = 1 & \lim_{x\to 2^-} g(x) = 1 \\ \lim_{x\to 1} g(x) \text{ does not exist.} & \lim_{x\to 2} g(x) \text{ does not exist.} \end{array}$$

So the two-sides limit $\lim_{x\to c} g(x)$ does not exist for $c = 1$ or $c = 2$. It does exist for all other values of c between 1 and 3 however.

Note, as usual, that the actual *values* $f(1) = 1, f(2) = 1, g(1) = 1, g(2) = 2$ do not matter for the existence of the limit, since the limit is idealizing experiments which do not hit the value exactly.

(b) For all value of c, $\lim_{x\to c^-}(f(x) + g(x)) = \lim_{x\to c^-} f(x) + \lim_{x\to c^-} g(x)$ (assuming that both these limits exist), and similarly for limits from the right. For all values of c besides 1 and 2, this shows that the two one-sides limits of $f(x) + g(x)$ exist and are equal. Now look what happens at $c = 1$ and $c = 2$.

$$\begin{aligned} \lim_{x\to 1^-}(f(x) + g(x)) &= \lim_{x\to 1^-} f(x) + \lim_{x\to 1^-} g(x) \\ &= 1 + 0 \\ &= 1 \\ \lim_{x\to 1^+}(f(x) + g(x)) &= \lim_{x\to 1^+} f(x) + \lim_{x\to 1^+} g(x) \\ &= 0 + 1 \\ &= 1 \end{aligned}$$

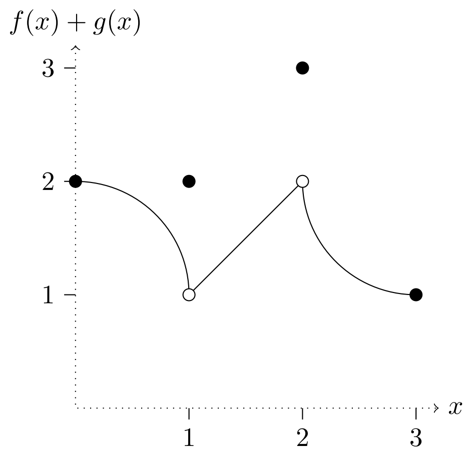
So despite the fact that the limit exists at $c = 1$ for neither $f(x)$ nor $g(x)$, in fact the limit of the sum does exist: $\lim_{x\to 1}(f(x) + g(x)) = 1$. What has happened is that their two "jumps" cancel out.

In fact, the same thing happens at $c = 2$.

$$\begin{aligned} \lim_{x\to 2^-}(f(x) + g(x)) &= \lim_{x\to 2^-} f(x) + \lim_{x\to 2^-} g(x) \\ &= 0 + 2 \\ &= 2 \\ \lim_{x\to 2^+}(f(x) + g(x)) &= \lim_{x\to 2^+} f(x) + \lim_{x\to 2^+} g(x) \\ &= 1 + 1 \\ &= 2 \end{aligned}$$

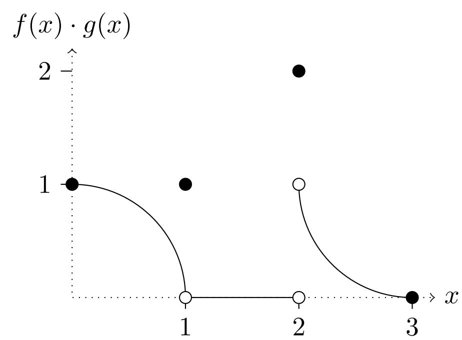
So the limit also exists at 2: $\lim_{x\to 2}(f(x) + g(x)) = 2$.

So in fact, the **limit exists at all values** c **between** 1 **and** 3. In fact, if you draw the graph of $f(x) + g(x)$, this is what it looks like.



This graph is very strange: although the limit exists at every point, the limit is not always equal to the value of the function. I encourage you to think through why the graph looks this way.

(c) Probably the easiest way to draw this graph is to notice that multiplying by $f(x)$ simply leave $g(x)$ the same (since they are multiplied by 1) at all values except those in the interval $(1, 2)$, which are sent to 0. The graph looks as follows.



Notice that the limit of $f(x)g(x)$ exists at all points except $x = 2$, even though the limit is not equal to the value of the function at $x = 1$.

Example 4.3. What is the slope of the tangent line through $(4, 2)$ on the graph of $y = \sqrt{x}$?

Solution 1. Suppose that $x \neq 4$, and consider the secant line through $(4, 2)$ and $(x, \sqrt{x})$. The slope of this secant line is given by the following function.

$$s(x) = \frac{\text{rise}}{\text{run}} = \frac{\sqrt{x} - 2}{x - 4}.$$

If you compute this function for some values of x close to $x = 4$, you will see for example that $s(5) \approx \frac{0.24}{1} = 0.24$, $s(4.1) \approx \frac{0.0248}{0.1} = 0.248$, and $s(4.01) \approx \frac{0.002498}{0.001} = 0.2498$. So these value seem to tending to 0.25, but as usually we cannot actually perform the ideal experiment and plug in $s(4)$.

In this case, there is a little algebraic trick to compute the limit: we can rationalize the numerator. You may have seen this trick in a different form, namely rationalizing the denominator to simplify expressions; here we are essentially doing the reverse. Observe that for all x,

$$\begin{aligned}
\frac{\sqrt{x}-2}{x-2} &= \frac{\sqrt{x}-2}{x-2} \cdot \frac{\sqrt{x}+2}{\sqrt{x}+2} \\
&= \frac{x-2}{(x-2)(\sqrt{x}+2} \\
&= \frac{1}{\sqrt{x}+2}
\end{aligned}$$

Actually, the last step only works for $x \neq 2$, since the terms we cancel on the top and bottom would be 0 at $x = 2$. Nevertheless, this equation hold at all nearby values of x, and we can therefore see that the ideal value of $s(x)$ for x near 4, that is the limit $\lim_{x \to 4} s(x)$ is equal to $\frac{1}{\sqrt{4}+2} = \frac{1}{4} = 0.250$.

Solution 2. For a slightly less standard solution, observe that $y = \sqrt{x}$ is the same as $y^2 = x$ (for y positive, at least). So we can consider the slopes of secant lines from $(4, 2)$ to (y^2, y), for values of y very close to 2. The slope of such a secant line is $\frac{y-2}{y^2-4} = \frac{y-2}{(y+2)(y-2)} = \frac{1}{y+2}$, so the limit as y goes to 2 is $\frac{1}{4}$, as before.
