One-variable Calculus — Limits and continuity

Welcome to Calculus! Here are my supplemental notes for one-variable Calculus, giving alternative ways to think about some things, practical advice, and sometimes more theoretical detail.

This does not cover everything that you need to know (although it covers a lot); you should also have the official course textbook, which is the 3rd Edition of *University Calculus: Early Transcendentals* by Hass et al published by Addison–Wesley (Pearson). There are also some references in these notes to that textbook.

Preliminaries

Before beginning this class, you should be familiar with the basic algebraic properties of real numbers.

By default, all of the numbers that we work with will be real numbers; most of Calculus applies just as well to complex numbers, but a complete understanding of Calclus in even one complex variable requires some ideas from multivariable Calculus, which these notes do not cover. In particular, if a is a negative number, then $\sqrt[n]{a}$ is undefined when n is an even integer and negative when n is an odd integer. More generally, if a is a negative number, then a^p is defined *only* if p is a rational number whose denominator in lowest terms is odd; in this case, a^p is positive if the numerator of p is even and negative if the numerator of p is odd. Note that $(a^2)^{1/2} = \sqrt{a^2} = |a|$, while $a^{2\cdot 1/2} = a^1 = a$, which is different when a is negative, so the rule that $(a^x)^y = a^{xy}$ does not hold in general (although it does hold when a is a positive number).

Although 0^x is undefined whenever x is negative (because this amounts to dividing by zero), we need to define $0^0 = 1$ in order to make some formulas work correctly. Although the course textbook says that 0^0 is undefined, this contradicts some things that the textbook says about polynomials and power series. (Section 9.7 of the textbook, beginning with the definition of power series on page 523, is the first place where this is important; see also the discussion of power series in the full version of these notes.) It's possible to take a more nuanced approach where 0^x is 1 when x is a *constant* with the value 0 while 0^x is undefined when x is a *variable* with the value 0; however, this makes the meaning of 0^0 ambiguous without context, so I prefer to simply say that $0^0 = 1$. Nevertheless, this will require some care when it comes to rules for evaluating limits.

The main difference in my approach to Calculus from the textbook's is that I make more use of differentials. Calculus was originally developed using differentials, and many calculations are easier to do this way. Furthermore, differentials are often used in applications, especially (but not only) to physics. They fell out of fashion with mathematicians towards the end of the 19th century, when Calculus was first put on a rigorous logical foundation, because this foundation did not include differentials. However, that problem was solved by the middle of the 20th century, so there is no longer any reason to avoid them. You can do almost everything in the textbook's way if you want, but I encourage you to try using differentials.

Functions

Another difference between these notes and the textbook is that I will never be sloppy with function notation.

In an expression such as

$$y = f(x),$$

the variables x and y stand for real numbers, while the variable f stands for a function. (Usually this variable is actually a constant, because f always refers to the same function throughout the problem, although there can also be situations where the function itself is allowed to vary.) A function is not a number but rather a process for turning one number into another. When speaking of specific numbers, this is usually not a problem; for example, $f(2) = 4$ means that the function f is a process that (among other things) turns the number 2 into the number 4.

The statement that $f(x) = x^2$ is more ambiguous; in a context where the variable x already appears, this means that the function f is a process that turns the number x (whatever number that is) into the number x^2. But in a context where x does not already have a meaning, this statement usually means that the function f is a process that turns *every* real number into its square, which is a complete description of the function. In this case, it is better to say something like

$f(x) = x^2$ for all x,

and I will usually say something like this.

Another way to completely describe this function is to write

$$f = (x \mapsto x^2).$$

This is analogous to defining a set S as $S = \{x \mid x > 2\}$; in each case, you introduce a new *dummy variable* and then give an expression (to define a function) or an equation or inequality (or other statement, to define a set) that refers to that dummy variable. You can even do this without giving the function (or set) a name, by just referring to the function $(x \mapsto x^2)$. Although the textbook does this with sets, it never does this with functions; so in order to avoid burdening you with additional notation to learn, I will not do this either. It can be very handy, however.

The real problem is when the same symbol is used both to refer to a function and to its output value, as in

$$A = A(x),$$

which you might see (for example) in a problem in which the area of some shape depends on something else. I will never do this! Either I will use A to refer to the area itself, or I will use A to refer to the function that indicates how this area depends on x, but I will not use the same symbol for both of these. If I need to refer to both of these, then I will use two different symbols. Most of the time, however, it's enough to have a symbol for the area itself and to leave the function unnamed. (The notation for evaluation described below on page 3 can help with this.)

When we cover derivatives later on, there are various symbols used for this concept; and when $y = f(x)$, then I will also write

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f'(x).$$

(What this means is explained in the full version of these notes.) The textbook will sometimes write y' or $\mathrm{d}f/\mathrm{d}x$ in this situation, but I never will, and this is important to ensure that the ordinary rules of algebra continue to apply to such expressions. (For example, you can multiply both sides of the equation above by $\mathrm{d}x$ to get $\mathrm{d}y = f'(x)\,\mathrm{d}x$, which would be difficult to do correctly using the wrong symbols.) I will not count it against you if you are as sloppy as the textbook about this, because I don't think that it's fair to require more of you than the textbook writers can manage; however, if you get confused by your notation and make a mistake, then that will count against you! So I encourage you to follow my notation.

Variables

In Calculus, we study *variable* quantities, that is quantities whose values may vary (or change).

In Algebra, we often use the word 'variable' to refer to any quantity whose value we don't know, even if this value is fixed and never changes throughout the problem. In fact, the standard Algebra problem, solving an equation such as $2x + 3 = 5$, involves figuring out the value of the variable; so it had only one value all along, and we just had to figure out what it was. So if x is a variable in an Algebra problem, and at some point we decide that the value of x is 1, then this may well mean that x is 1 throughout the entire problem. (That's not always the case in Algebra, but it often is.)

In Calculus, we take the word 'variable' more seriously. If x is a variable in a Calculus problem, then x might be 1 at some point, but it will probably be 6 at some other point in the problem. (And more often than not, it will take all of the values in between 1 and 6 along the way, such as $1\frac{1}{2}$, π, and 5.789.) Furthermore, if x and y are two variables that appear in the same problem, then the value of y will usually change as the value of x changes. Calculus is primarily about exactly this sort of thing: *how* one quantity changes as another quantity changes.

In the simplest cases, it turns out that y is a function of x; that is, there is a fixed function f such that $y = f(x)$ remains true as x and y vary. Calculus textbooks generally try to fit everything into this mould, but it doesn't always come out like this naturally. Often, you know that both x and y are changing, but it's not obvious that the value of x at some point is enough information to figure out the value of y at that point; yet when you write $y = f(x)$, you're assuming that it is enough information.

Most of the time, however, we can assume that there is some variable t, called the *independent variable*, such that every other variable in the problem is a function of t. That is, if x and y appear in the problem, then there are fixed functions f and g such that $x = f(t)$ and $y = g(t)$ throughout the problem. (Then x and y are called *dependent variables*, since their values depend on the values of t, through the functions f and g.) But this variable t might not show up directly! Calculus books will usually tell you (especially in word problems) that it's necessary to pick an independent variable, but it's enough to visualize the range of variation in the problem, and you can treat all of the variables on an equal footing. All the same, for the sake of formal definitions, we will assume that there is an independent variable t and that every other variable is a function of it, even though in practice we don't have to identify it. (Of course, you don't have to call the independent variable 't', but I usually will, just to have a consistent name.)

If we're not going to refer directly to t, then we're not going to refer directly to f and g either, so we need some way to refer to the values of these functions without referring to the functions themselves. Here is how we do it formally:

If $u = f(t)$, then $u|_{t=c} = f(c)$.

(This is called *evaluation notation*.) More generally, if P is some statement that is only true once, then P is implies the statement $t = c$ for some value of c, so we can make sense of $u|_P$. Even if P is a statement that might not only be true once, as long as every possible value of $u|_P$ is the same, then we can still make sense of $u|_P$. Finally, even if there are different possible values of $u|_P$, then the value of $u|_P$ still varies, but at least it doesn't vary as much as u itself, since there are now fewer possibilities.

This all sounds very abstract (because it is), but the concrete application is straightforward; here are some examples:

$$\begin{aligned} x|_{x=5} &= 5, \\ (2x+3)|_{x=4} &= 2(4)+3 = 11, \\ (2x+3y)|_{\substack{x=4,\\ y=5}} &= 2(4)+3(5) = 23. \end{aligned}$$

Taking the last of these for example, there is no need to think about what t is when $x = 4$ and $y = 5$, and indeed without considering how x and y depend on this unspecified independent variable t, the value of t is impossible to know. Nevertheless, we know that no matter what t may be, if $x = 4$ and $y = 5$, then $u = 2x + 3y$ is definitely $2(4) + 3(5) = 23$, and that is enough. So all that you have to do in practice is to plug and chug.

Sometimes (generally only in the middle of a problem or in something theoretical) you can't work out the value completely; for example,

$$(2x+3y)|_{x=4} = 2(4) + 3(y|_{x=4}) = 8 + 3y|_{x=4}.$$

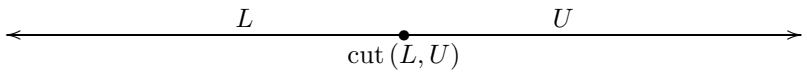
If we don't know anything more about the relationship between x and y, then we don't know the value of y when $x = 4$, so this is all that we can say in this example, but at least we were able to work out part of it.

Real numbers

In this course, we work with the real numbers, which are supposed to correspond to points on a number line. Ultimately, all of the properties of real numbers derive from intuitive geometric properties of points on a line. For example, the arithmetic operations of addition, subtraction, multiplication, and division can be defined in terms of changes of position and scale on the number line. The order relation between real numbers ($<$ and $>$) also derives from relative position on a line. (You have to specify the numerical values of at least two points, such as 0 and 1, in order to make a geometric line into a number line, but once you have those two points, then everything else follows.)

The most advanced of the fundamental properties of the number line is its *completeness*. There are many ways to express completeness, but my favourite is this:

If you pick out two nonempty regions of the number line, one on the left called L and one on the right called U, which don't overlap but otherwise cannot be extended further, then there is a single point between them, called $\text{cut}\,(L,U)$, the *cut* between L and U.



We can make this logically precise (in terms of the order relation on real numbers): Suppose that L and U are two sets of real numbers (making precise what regions of the number line are), with these properties:

- There is some $r \in L$ and some $s \in U$ (which is what it means for L and U to be nonempty);
- If $r \in L$ and $s \in U$, then $r < s$ (which is what it means for L to be on the left and U on the right without overlapping);
- If $r < s$, then $r \in L$ or $s \in U$ (which is what it means to say that L and U cannot be extended further).

(Note that 'or' in math, as here, normally includes the possibility of both.) Then there exists a real number $\text{cut}\,(L,U)$ with this property:

- If $r \in L$ and $s \in U$, then $r \le \text{cut}\,(L,U) \le s$ (which is what it means for $\text{cut}\,(L,U)$ to be between L and U).

A couple more important properties follow from what was said above:

- The number $\text{cut}\,(L,U)$ is the *only* real number between L and U;
- If $r < \text{cut}\,(L,U) < s$, then $r \in L$ and $s \in U$.

The point of all this is to be able to prove that a real number exists. For example, in order to prove rigorously that every real number c has a cube root $\sqrt[3]{c}$ (and has anybody ever showed you why this is true or did you just take it on faith?), you first define L as $\{x \mid x^3 < c\}$ and U as $\{x \mid x^3 > c\}$, check that L and U have the necessary properties listed above, conclude that $\text{cut}\,(L,U)$ exists, and check (using the properties of $\text{cut}\,(L,U)$ listed above) that $\text{cut}\,(L,U)^3 = c$. Thus, this cut is the cube root $\sqrt[3]{c}$.

This method of proving that a real number exists is also practical, because it shows us how to approximate its value as closely as we like. For example, to approximate $\sqrt[3]{2}$ to 4 decimal places, you look at some nearby possibilities, such as $1.0001, 1.0002, 1.0003, \ldots, 1.9997, 1.9998, 1.9999$. Somewhere in this list are two numbers right next to each other, one of which has a cube less than 2 (so it's in L) and one of which has a cube greater than 2 (so it's in U). Then we approximate $\sqrt[3]{2}$ to 4 decimal places by saying that it's in between these two numbers. (As it happens, these two numbers are 1.2599 and 1.2600; also, $1.25995^3 > 2$, so $\sqrt[3]{2}$ rounds to 1.2599.) There are more efficient ways to calculate cube roots, but this proof that they exist at least gives *one* way to calculate them, to start with.

Limits and continuity

There are four main operations considered in Calculus: limits, derivatives (or differentials), integrals (or antidifferentials), and sums of infinite series. (The last of these is not covered until next term.) Here we will look at the first one: limits. These are also closely related to the concept of continuity, which is actually the easiest concept to define.

Continuity

In Calculus, we not only study variable quantities; we study quantities that are *continuously* varying. This implies in particular that a quantity y that varies from 1 to 6 will pass through $1\frac{1}{2}$, π, and 5.789, and everything else in between.

In real life, we can never measure or fix the value of a such a quantity y exactly, down to the last decimal place; after all, there are infinitely many decimal places, but we can only do a finite amount of work. So, it is key to the study of real numbers that we can *approximate* them to any finite number of decimal places (among other ways). That is what the stuff about cuts above accomplishes.

Also in Calculus, we study how one quantity y varies along with another quantity x. This may be expressed by saying that y is a *function* of x; if f is the function, then $y = f(x)$. But in practice, we only know x and y *approximately*, so if we only use an approximate value of x, then $f(x)$ should still be an approximate value of y. For example, suppose that $f(x) = x^2$; if you know that x is approximately 2, then you know that $y = f(x)$ is approximately $2^2 = 4$.

This doesn't work with every function! For example, suppose that g is the piecewise-defined function

$$g(x) = \begin{cases} x+1 & \text{for } x < 2, \\ x+3 & \text{for } x \geq 2; \end{cases}$$

if you only know that x is approximately 2, then you really don't know if $g(x)$ is approximately $2+1=3$ or approximately $2+3=5$. Of course, if you knew that x is *exactly* 2, then you would know that $g(x)$ is $2+3=5$; but it's no good if you only know x approximately.

In these examples, we say that g has a **discontinuity** at 2, while f is **continuous** at 2. (In fact, f is continuous everywhere, while g is continuous everywhere except at 2.) So the idea is this:

> A function f is **continuous** at a real number c if, whenever $x \approx c$ (meaning that x is approximately equal to c), $f(x) \approx f(c)$.

So if you only know that $x \approx c$, then that's enough information to know $f(x)$ approximately (specifically, that $f(x) \approx f(c)$).

Actually, we should take care about where f is defined. Sometimes Calculus textbooks say that f has a discontinuity at c if $f(c)$ is undefined, and sometimes they don't; but in any case, f is not continuous there: f must be defined first in order to be continuous. On the other hand, if $f(x)$ is undefined, then we don't hold that against f; for example, we want to say that $f(x) = \sqrt{x}$ is continuous at 0, even though $f(x)$ is undefined (as a real number) whenever $x < 0$. So a more careful definition is this:

> A function f is **continuous** at a real number c if $f(c)$ is defined and, whenever $x \approx c$ and $f(x)$ is defined, $f(x) \approx f(c)$.

This is still not a completely rigorous definition, because it doesn't explain how close we need to be to say that one quantity is approximately equal to another. (Basically, the answer is this: as close as you need, and as close as you want.) But I will save that for a bit later. Already, this basic idea should be enough to allow you to judge continuity of a function from its graph.

To judge continuity of a function from a formula, know that any function is continuous (wherever it is defined) if it has a formula that uses only these operations: addition, subtraction, multiplication, division, absolute values, exponentation if the exponent is constant or the base is always positive, roots if the index is constant or the radicand is always positive, logarithms, trigonometric functions, and inverse trigonometric functions. These are pretty much all of the functions that you ever deal with!

So, the exceptions are much rarer: exponentiation where the exponent varies and the base can be zero or negative, roots where the index varies and the radicand can be zero or negative, and piecewise-defined functions. Of these, only piecewise-defined functions are likely to come up. These functions *can* be continuous, but only if the values agree on both sides whenever two pieces join. So for example, while

$$g(x) = \begin{cases} x+1 & \text{for } x < 2, \\ x+3 & \text{for } x \geq 2 \end{cases}$$

has a discontinuity at $x = 2$,

$$h(x) = \begin{cases} x+1 & \text{for } x < 2, \\ 5-x & \text{for } x \geq 2 \end{cases}$$

is continuous at $x = 2$ (and so everywhere), because $2+1 = 5-2$.

Returning to the meaning of continuity, how close of an approximation is close enough? The key to the answer is that a real number may be approximated as precisely as you wish, as long as you put enough work into it. So for f to be continuous at c, we should be able to demand that $f(x)$ and $f(c)$ be as close together as we like (as long as we still allow for a positive distance between them). But in order to achieve that result, it's fair in turn to demand that x be as close to c as necessary (again as long as we still allow the distance to be positive). The distance between two numbers is given by subtracting and taking the absolute value, so we need to be able to ensure that $|f(x) - f(c)|$ is as small as we want (but positive) by making $|x - c|$ as small as we need (but positive).

The traditional symbols for these small but positive distances are the Greek letters 'ϵ' (Epsilon) and 'δ' (Delta). For this reason, this is sometimes called the ϵ-δ (or epsilon-delta) definition; this method is also called *epsilontics*. So here is the rigorous definition:

A function f is **continuous** at a real number c if $f(c)$ is defined and, for each positive number ϵ (no matter how small), there is some positive number δ (possibly quite small), such that whenever $|x - c| < \delta$ and $f(x)$ is defined, $|f(x) - f(c)| < \epsilon$.

This is fairly complicated, but you can view it as a game, involving a function f and a number c such that $f(c)$ exists.

- I challenge you with a positive number ϵ.
- You respond with a positive number δ.
- I reply with a value of x such that $|x - c| < \delta$ and $f(x)$ is defined.
- You win if $|f(x) - f(c)| < \epsilon$.

If you can win this game, no matter what choices I make, then f is continuous at c. On the other hand, if I can win no matter what choices you make, then f has a discontinuity at c.

Directions and limits

A **direction** in some variable describes not only whether the variable is increasing or decreasing (that is its literal direction on a number line) but also if there is a limiting value that it approaches but does not reach. The basic directions that we study in this course take the following four forms, where x may be any variable and c may be any constant:

- as x increases without bound: $x \to \infty$;
- as x decreases without bound: $x \to -\infty$;
- as x increases towards c: $x \to c^-$;
- as x decreases towards c: $x \to c^+$.

Any two or more of these directions may be combined, but the only type of combined direction in the textbook is this:

- as x approaches c: $x \to c$;

which is the combination of $x \to c^-$ and $x \to c^+$. That said, other combinations are also sometimes studied, especially the combination of $x \to \infty$ and $x \to -\infty$, which is written $x \to \pm\infty$. (You can also consider fancier directions, for example as x increases without bound *while taking only integer values*, which is relevant to the material in Section 9.1 of the textbook. For now, however, I'll stick to the 5 types of directions covered in Chapter 2.)

If D is any direction and u is any variable quantity, then we indicate the value to which u approaches as change occurs in the indicated direction as

$$\lim_{D} u$$

in a diplayed equation or as $\lim_D u$ in running text. (The textbook likes to write u as $f(x)$, and this is certainly convenient when it comes to the formal definition, but in practice you'll start with an expression involving the variable x, and it's not necessary to think of this as given by a function.) We will examine the case when u approaches a real value L, as well as the case when u increases without bound or decreases without bound. In the first case, we say that the limit **converges**; in the second case, we say that the limit **diverges** to (positive or negative) infinity. (Other types of behaviour are also possible, which are also kinds of divergence, but for now I'll stick to the 3 types of results covered in the textbook.) A limit as x approaches c is one of the three kinds of results that we are considering if and only if the limits as x increases and decreases towards c are both this same result.

So in total, there are fifteen kinds of limits that we will consider, for the five kinds of directions (four basic and one combined) and the three kinds of answers:

$$\begin{array}{lll}
\lim\limits_{x\to\infty} u = L; & \lim\limits_{x\to\infty} u = \infty; & \lim\limits_{x\to\infty} u = -\infty; \\
\lim\limits_{x\to-\infty} u = L; & \lim\limits_{x\to-\infty} u = \infty; & \lim\limits_{x\to-\infty} u = -\infty; \\
\lim\limits_{x\to c^-} u = L; & \lim\limits_{x\to c^-} u = \infty; & \lim\limits_{x\to c^-} u = -\infty; \\
\lim\limits_{x\to c^+} u = L; & \lim\limits_{x\to c^+} u = \infty; & \lim\limits_{x\to c^+} u = -\infty; \\
\lim\limits_{x\to c} u = L; & \lim\limits_{x\to c} u = \infty; & \lim\limits_{x\to c} u = -\infty.
\end{array}$$

To see how to read these aloud, I'll consider the last one as an example; this says that the **limit**, as x approaches c, of u is negative infinity.

It's sometimes convenient to think of ∞ and $-\infty$ as numbers like the real number c or L, only numbers of an infinite magnitude. Similarly, it's sometimes convenient to think of c^+ and c^- as numbers that are infinitely close to (but distinct from) the real number c. Then the meanings of the directions are as follows:

- $x \to \infty$: what happens when x is positive and infinite?
- $x \to -\infty$: what happens when x is negative and infinite?
- $x \to c^-$: what happens when x is infinitely close to but less than c?
- $x \to c^+$: what happens when x is infinitely close to but greater than c?
- $x \to c$: what happens when x is infinitely close to but distinct from c?

Similarly, the meanings of the results are as follows:

- $\lim_D u = L$: u is equal to or infinitely close to L;
- $\lim_D u = \infty$: u is positive and infinite;
- $\lim_D u = -\infty$: u is negative and infinite.

This can be made into a rigorous definition, by extending the real number system to the *hyperreal* number system, although this is not the basis of the definition that we will be using. Still, it can be useful for intuition.

There are some other notations for limits that are sometimes useful. First of all, instead of writing $\lim_D u$, you can also write $u|_D$, analogous to evaluation notation (page 3). That is, $u||_{x=c}$ means the value that u *is* when x *is* c, while $u||_{x\to c}$ means the value that u *approaches* when x *approaches* c.

The point of a continuous function is that these are the same; that is, f is continuous at c if and only if $f(x)|_{x=c}$ and $f(x)|_{x\to c}$ both exist and are equal. Of course, instead of writing $f(x)|_{x=c}$, you could just write $f(c)$; similarly, instead of writing $f(x)|_{x\to c}$, there is yet another notation for this:

$$f(c^\pm) = f(x)|x \to c = \lim_{x\to c} f(x).$$

You can read this as 'f of c plus or minus'; the idea behind 'plus or minus' here is the same as in the English phrase 'more or less', meaning 'approximately', because we're looking at values of f *near* c rather than *at* c. Then f is continuous at c if and only if $f(c^\pm) = f(c)$.

The analogous notations for the other types of directions are $f(c^-)$, $f(c^+)$, $f(\infty)$, and $f(-\infty)$. Since things like c^- and ∞ aren't real numbers, there should be no confusion between this function-limit notation and the usual function-evaluation notation $f(c)$. Since all of these alternative notations for limits aren't in the textbook, I won't use them, but they are good to know; they are short and handy, and you may see them elsewhere.

Defining limits

The simplest type of limit to define is $\lim_{x\to c} f(x)$. Note that this just depends on the function f and the real number c, which is especially clear using the notation $f(c^\pm)$. If f is continuous at c, then this is supposed to be $f(c)$. But what if f is undefined or discontinuous at c?

Given a real number L, let $f_{c\mapsto L}$ be the piecewise-defined function given by

$$f_{c\mapsto L}(x) = \begin{cases} f(x) & \text{if } x \neq c, \\ L & \text{if } x = c. \end{cases}$$

That is, $f_{c\mapsto L}$ is almost the same function as f, except that $f_{c\mapsto L}(c) = L$, regardless of what $f(c)$ is (or even whether $f(c)$ is defined in the first place). Now here is the definition of the limit:

> If there is a unique real number L such that $f_{c\mapsto L}$ is continuous at c, then L is the limit of f approaching c.

Note that the limit is undefined if either there is no L that makes $f_{c\mapsto L}$ continuous or if there is more than one L that makes it continuous. But that second possibility is very rare; it only happens if f is not defined anywhere near c (in which case $f_{c\mapsto L}$ is continuous at c no matter what L is, because there is nothing nearby to compare to.)

What if the limit is some kind of infinity? We can't talk about $f_{c\mapsto\infty}$, because then $f_{c\mapsto\infty}(c)$ would have to be ∞, which is not a real number. However, if $f(x)$ is increasing without bound, then $1/f(x)$ should be approaching 0. This *almost* allows us to define the limit as $x \to \infty$; the only problem is that $1/f(x)$ still approaches 0 even if $f(x)$ *decreases* without bound as well. To finish the definition, therefore, we need a couple more functions (which are also useful in other contexts):

$$f^+(x) = f(x) \text{ for } f(x) \geq 0,$$
$$f^-(x) = f(x) \text{ for } f(x) \geq 0.$$

These are partially-defined functions; the first is undefined whenever its value would have been negative, while the second is undefined whenever its value would have been positive. Using these will force us to distinguish between when the limit is ∞ (positive values) and when the limit is $-\infty$ (negative values).

Now here are the definitions:

$$\lim_{x\to c} f(x) = \infty \;\Leftrightarrow\; \lim_{x\to c}\left(\frac{1}{f^+(x)}\right) = 0,$$
$$\lim_{x\to c} f(x) = \infty \;\Leftrightarrow\; \lim_{x\to c}\left(\frac{1}{f^-(x)}\right) = 0,$$

where the symbol '$\Leftrightarrow$' means 'if and only if' (that is, the statement on either side is true whenever the statement on the other side is true).

Finally, for the general defintion of $\lim_D u$, where D is any direction and u is any expression, suppose (like I did back on the bottom of page 2) that x and u are both functions of some independent variable t, where x is the variable that appears in the direction D. To be precise, suppose that $u = f(t)$ and $x = g(t)$. If the direction D consists of some additional condition on the variable x, then assume that this condition holds for every value of the function g. (So for $x \to c^-$, suppose that $g(t) < c$ always, and for $x \to c^+$, suppose that $g(t) > c$ always.) Then if the limit of $f(t)$ has the same value (a real number L, ∞, or $-\infty$) whenever the limit of $g(t)$ is the value given by the direction D (a real number c, ∞, or $-\infty$), then that value for the limit of $u = f(t)$ is the limit $\lim_D u$.

The official textbook defines limits directly using epsilontics (which is very similar to the epsilontic definition of continuity but slightly more complicated), then defines continuity using limits; I have defined continuity using epsilontics and defined limits using continuity. Our definitions come in different orders, but they are equivalent. In any case, the most important method of calculating limits is this:

- If f is continuous at c, then $\lim_{x\to c} f(x) = f(c)$.

This fact makes *most* limits trivial to calculate; but it's the exceptions where all of the interesting stuff happens!

More to come!