Review of calculus: part 1

This brief review recalls some of the most important concepts, definitions, and theorems from basic calculus. It is not intended to teach basic calculus from scratch. If any of the items below is not familiar to you, please consult a calculus book and review the topic in more detail.

Limits and continuity

Calculus begins with the **definition of limit**. In fact, there are two, related definitions. We say that a sequence of numbers a_n approaches the limit L in case for all positive ϵ there exists a real number N such that $|a_n - L| < \epsilon$ whenever $n > N$. This is the definition of **sequential limit**. We write $\lim_{n\to\infty} a_n = L$ as shorthand notation to mean that the sequence a_n approaches the limit L. A second notation for this is to write $a_n \to L$ as $n \to \infty$.

The second, related definition is **functional limit**: we say that $f(x)$ approaches L as x approaches a in case for all positive ϵ there exists a positive δ such that $|f(x) - L| < \epsilon$ whenever $0 < |x - a| < \delta$. As above, we use the notation $\lim_{x\to a} f(x) = L$ and also the notation $f(x) \to L$ as $x \to a$.

One can also define **one-sided limits**: for the left-hand limit we change the definition to require that $|f(x) - L| < \epsilon$ when $a - \delta < x < a$; for the right-hand limit the requirement holds when $a < x < a + \delta$. The left- and right-hand limits are denoted $\lim_{x\to a^-} f(x)$ and $\lim_{x\to a^+} f(x)$, respectively.

It is a theorem that $\lim_{x\to a} f(x) = L$ if and only if whenever $a_n \to a$ we have that $f(a_n) \to L$. It is also a theorem that f has a limit at a if and only if f has both a left- and a right-hand limit and these limits are equal.

The notion of limit is the basis for the definition of **continuity**: we say that f is continuous at a if $\lim_{x\to a} f(x) = f(a)$. Said another way, f is continuous at a if the following three conditions are satisfied:

- f is defined at a.
- f has a limit at a.
- The limit equals $f(a)$.

We can interpret limits in the context of approximations. To say that $f(x) \to L$ as $x \to a$ is to say that $f(x) \approx L$ when $x \approx a$. More precisely, to say that $|f(x) - L| < \epsilon$ is to say that $L - \epsilon < f(x) < L + \epsilon$, which is to say that $f(x) \approx L$ with error less than ϵ.

There are many ways a function can be discontinuous. One way — the most important for us — is to have a **removable discontinuity**. This happens when f has a limit L at a point a, but either f is not defined at a or else $f(a) \neq L$. In this case, if we redefine $f(a)$ to be L, then the new definition gives a function which is continuous at a.

If both the left- and right-hand limits exits, but are not equal to one another, then we say that f has a **jump discontinuity**. Some of the most important examples of functions with a jump discontinuity are the **unit step functions** — sometimes called the **Heaviside functions** — which are defined by the rule

$$u_c(x) = \begin{cases} 0, & \text{if } x < c \\ 1, & \text{if } x \geq c \end{cases}$$

A third type of discontinuity is a vertical asymptote. This happens when for all M there is a positive δ such that $|f(x)| > M$ whenever $0 < |x - a| < \delta$. We sometimes use the notation $\lim_{x\to a} f(x) = \infty$ or $f(x) \to \infty$ as $x \to a$.

There are many, many other kinds of discontinuities. Just to give yourself an inkling of how bad a discontinuity can be, look at a graph of the function $\sin(1/x)$ near 0.

One of the main properties of a continuous function is the **Intermediate Value Theorem**, which says that if f is continuous on a interval containing points a and b, and if z is any value between $f(a)$ and $f(b)$, then there is a point c between a and b such that $f(c) = z$. This is often used in the following special case: if f is continuous on some interval, and if f takes a positive value somewhere on the interval and a negative value at another, then somewhere in between f must equal 0.

The Intermediate Value Theorem leads to a method for solving equations defined by continuous functions. If f is continuous on the interval $[a, b]$ and if f changes sign between a and b — that is, $f(a)$ and $f(b)$ are of opposite sign — then either f changes sign somewhere between a and the midpoint $(a+b)/2$, or else between this midpoint and b. In the former case we replace b by the midpoint; in the latter we replace a by the midpoint. In either case we have replaced the interval by one half as long, and so have narrowed by half the possible error in knowing where $f(x) = 0$. This is called the **Bisection method**. It is very slow, but has the advantage that it includes a easily predictable error estimate.

One of the main theorems is the **continuity of algebra**, which says that any function defined by algebraic operations — addition, subtraction, multiplication, and division — is continuous at every point of the domain (that is, at any point where you don't try to divide by 0). In particular, all polynomial functions are continuous everywhere.

It is also a theorem that if g is the inverse function of a one-to-one function f, and if f is continuous at a, then g is continuous at $f(a)$. In particular, the fractional powers $x^{1/n}$ are all continuous wherever they are defined.

Slope and derivative

A nonvertical line has constant **slope**. In other words, if points (a, b) is any point of the line and if $(a + \Delta x, b + \Delta y)$ is any second point, then the ratio $\Delta y/\Delta x$ always yields the same value, no matter what two points we choose. (Of course, when we change the line then the value of this slope changes.) The slope is the most important characteristic of the line. A positive slope means the line represents an increasing function; a negative slope corresponds to a decreasing function; and a constant function has slope 0.

Lines can be defined by many types of equations. For example $2x - 3y = 7$, $x = 9$, $y - 8 = 5(x + 4)$ are all equations which define lines. The last is an example of **point-slope formula**: if the line passes through the point (a, b) and has slope m then one equation for the line is $y - b = m(x - a)$. Every nonvertical line has a well-defined slope, and hence every point on it yields an equation in point-slope form. For us this will be the most useful formula for lines.

The **derivative** of f at a point a is defined to be

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x},$$

provided the limit exists. If it does we use either of the notations $f'(a)$ or $\frac{df}{dx}(a)$ to denote this limit, and refer to it as the **instantaneous slope** of f at a. When the derivative exists we say that the function is **differentiable**.

Another way to think about the derivative is to consider the slope function $m(\Delta x)$, which measures the slope of the secant line thru $(a, f(a))$ and $(a + \Delta x, f(a + \Delta x))$. This function is not defined when $\Delta x = 0$, but when it has a removable discontinuity there the value we define when $\Delta x = 0$ is the derivative of f at a.

A differentiable function can be approximated by a linear function on sufficiently small intervals. So, if you graph a differentiable function on a graphing calculator and then zoom in on small intervals then what you see is indistinguishable from a straight line. This is because if $\Delta y/\Delta x \to f'(a)$, then for any positive ϵ we have that $\Delta y \approx f'(a) \cdot \Delta x$, with error less than $\epsilon \cdot \Delta x$, whenever $|\Delta x|$ is sufficiently small.

One of the most important consequences of this interpretation is the theorem that if f is differentiable then f is continuous.

When we have a function of more than one variable, $f(x, y, \ldots)$ and if we hold all but one variable constant, and differentiate with respect to the remaining variable, we obtain the **partial derivative** of f. For example, if f is a function of x and y, then it has two partial derivatives, which we denote $\partial f/\partial x$ and $\partial f/\partial y$. We will make only slight use of partial derivatives this semester.

The higher order derivatives of a function are defined by iterating the operation of the differentiation. So, the **second derivative** is the derivative of the derivative, and is denoted either f'' or $\frac{d^2 f}{dx^2}$. The n-th derivative of f is denoted either $f^{(n)}$ or $\frac{d^n f}{dx^n}$.

The Mean Value Theorem

If f is continuous on a closed interval $[a, b]$ then there is a point c in that interval such that $f(c) \geq f(x)$ for every x in $[a, b]$. Such a value $f(c)$ is called the maximum of f on $[a, b]$. Note that there may be several points at which f takes its maximum value. Similarly, if f is continuous on $[a, b]$ the f takes a minimum value somewhere in $[a, b]$.

Fermat's Theorem is the observation that if the f takes on a maximum (or minimum) value at some point c in the interior (a, b), and if f is differentiable at c, then $f'(c) = 0$. When $f'(c) = 0$ we say that c is a **critical point** for f. So, if f has a maximum (or minimum) at c then either c is a critical point, an endpoint, or a point where f is not differentiable (sometimes called a **singularity**).

Using Fermat's Theorem we can prove the **Mean Value Theorem for Derivatives**, which says that if f is continuous on $[a, b]$, and differentiable on the interior (a, b), then there is a point c in the interior such that

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

The right-hand side of this equation is the average slope of f on $[a, b]$, whence the name of the theorem. We will state a generalization of this theorem to higher order derivatives when we come to power series.

The Mean Value Theorem has an important consequence for the study of differential equations — the **uniqueness of solutions to differential equations**. The simplest form of this says that of f and g are differentiable on (a, b), and if $f'(x) = g'(x)$ for all x in (a, b), then in fact $f(x) - g(x)$ is constant on the interval. In geometric language this says that if the graphs of f and g have the same (instantaneous) slope at every point, then the graph of one can be shifted up or down to match the graph of the other. A third way to state this theorem is to say that if we are given the value of the derivative of f at every point, and the value of f at one point, then there is at most one possibility for f.

When we have covered integration we will see in fact that there is always *exactly* one possibility for f, altho it may be difficult or even impossible to find a neat formula for it.

Homework problems: due Wednesday, 18 January

1 Compute the slope function $m(\Delta x)$ for the function x^2 at the point $(-2, 4)$, and sketch its graph. Show that x^2 is differentiable at -2 by observing that it has a removable discontinuity. What does the graph of $m(\Delta x)$ tell you is the value of the derivative x^2 at -2?

2 Compute the slope function $m(\Delta x)$ for the absolute value function at the point $(0, 0)$, and sketch its graph. Show that $|x|$ is *not* differentiable at 0 by showing that the slope function $m(\Delta x)$ does *not* have a removable discontinuity. What kind of discontinuity does it have?

3 Use the definition of derivative to compute $f'(a)$, for each of the following.

   a $f(x) = 1/x$, $a = 1/3$.

   b $f(x) = \sqrt{x}$, $a = 4$.

   c $f(x) = (x - 3)^9$, $a = 3$.

4 Use the Bisection Method to approximate the root of the equation $x^5 + x - 1 = 0$, accurate to within 0.01.

5 Let $f(x) = (x + 1)x^2 + (x - 3)x$. Find all points c in the interval $[-1, 3]$ at which the instantaneous slope equals the average slope over the interval.

6 Use Fermat's Theorem to find the maximum and minimum values of $|x| - x^3$ on $[-0.5, 1.5]$.