Math 131, Lecture 1

Charles Staats

26 September 2011

1 Introduction

Loosely speaking, there are two sides to mathematics: the ideas and the technical skills. Most people who say that they hate math have probably gotten hung up on the technical side. And it is an unfortunate fact that the technical side cannot be done away with. However, the ideas are what make the technical side interesting. Without them, no one would ever have discovered the technical side, and certainly no one would care to study mathematics as their life's work.

In this course, I will try to flavor the technical details with the ideas that explain why people were thinking like this in the first place. Think of studying mathematics like studying a map. One can simply sit down and try to memorize all the rivers, lakes, and mountain ranges. Or one can imagine how an explorer might travel, and bring the landforms to life. A river, for instance, becomes at once an obstacle, a water source, and a highway. Some rivers you can wade across; others are difficult enough that you may want to build a bridge. My goal is to present the mathematical "landforms" with some kind of narrative about how the first explorers might have seen them, and why they built the things they did.

The situation with calculus is especially tricky. The basics of calculus, as invented by Newton and Leibniz in the late 1600s, might be seen as "exploring on top of the clouds." There are plenty of interesting things to explore on top of these clouds, but you can never be sure what's under your feet. You might step on a spot that looks solid, only to find yourself standing on air. In the 1800s, mathematicians (most notably Cauchy and Weierstrass) built a solid "foundation" to fix this problem, which is not so much a foundation as a skyscraper. In this course, we will try to explore both the "castle in the clouds" and the skyscraper—called *analysis*—that holds it up. Even if you have seen some calculus before, you have almost certainly not seen much of the skyscraper.

2 Proofs

One of the key things that distinguishes mathematics from other disciplines is the presence of logical proofs. In physics, you know that a ball will fall when you release it because it has done so every time you released it in the past. But in some sense, there is no absolute reason why it would have to keep behaving in this fashion. One can imagine that gravity might suddenly stop working tomorrow.

In mathematics, we know that $\sqrt{2}$ is irrational because we can show, definitively, that it could never be rational. And in fact, the way we prove this is called *proof by contradiction*: imagine a world in which $\sqrt{2}$ were rational, and show that this world is contradictory; thus, it cannot exist.

The Fundamental Theorem of Arithmetic. Every natural number greater than 1 can be written as the product of primes in a unique way, except for the order of the factors.

For example, $45 = 3 \cdot 3 \cdot 5$.

Claim. Let n be a natural number greater than 1. Then n and n^2 have exactly the same prime factors.

Proof. If

$$n = p_1 \cdot p_2 \cdots p_k$$

is a factorization of n into primes $p_1, \ldots, p_k$, then

$$n^2 = p_1 p_1 \cdot p_2 p_2 \cdots p_k p_k$$

is a factorization of n^2 into primes. In both cases, the primes factors are $p_1, \ldots, p_k$. $\square$

It is a standard fact that any rational number can be expressed in *lowest terms*. That is, we can write it as a quotient p/q such that p and q have no prime factors in common.

Theorem. Let n be a natural number. Then $\sqrt{n}$ is either a natural number or an irrational number.

Proof. Assume, by way of contradiction, that $\sqrt{n}$ is rational, but is *not* a natural number. Writing $\sqrt{n}$ in lowest terms as

$$\sqrt{n} = \frac{p}{q},$$

where p and q are natural numbers with no prime factors in common. Since $\sqrt{n}$ is not a natural number, $q > 1$.

Claim. p^2/q^2 is already in lowest terms.

We want to see that p^2 and q^2 have no common prime factors. But the prime factors of p^2 are precisely the prime factors of p; likewise, the prime factors of q^2 are precisely the prime factors of q. And since p/q is in lowest terms, we know that p and q have no prime factors in common. This proves the claim.

Now, since $q > 1$ and both are positive, we know $q^2 > 1^2 = 1$. Hence, p^2/q^2 is not a natural number (the denominator is bigger than 1). But since $\sqrt{n} = p/q$, we can square both sides to find that

$$n = \frac{p^2}{q^2}.$$

Thus, n is not a natural number. Since we originally assumed it was, this is a contradiction. □

Corollary. $\sqrt{2}$ is irrational.

Proof. By the theorem above (with $n = 2$), $\sqrt{2}$ is either a natural number or an irrational number. Since $\sqrt{2}$ is not a natural number, it is irrational. □

Bonus Exercise. *Show that* $\sqrt[3]{31}$ *is irrational.*

Math 131, Lecture 2

Charles Staats

28 September 2011

1 Analysis is about inequalities, not equations

Traditional mathematics is about equations—determining when two quantities are equal. To "calculate" a quantity means, typically, to find an equal quantity that is easier to work with. For instance, when we convert a fraction to a decimal, we obtain the same number in a form that is easier to add to other numbers.

However, this notion breaks down when we try to deal with irrational numbers like $\sqrt{2}$. No matter how many digits of $\sqrt{2}$ we calculate, we will never find a decimal number equal to $\sqrt{2}$. The most we can do is to *approximate* $\sqrt{2}$. Thus, for instance, when we state that the first few digits of $\sqrt{2}$ are 1.414, we are really stating that

$$1.414 \leq \sqrt{2} \leq 1.415;$$

since all quantities are positive, we can square them to obtain the equivalent inequality

$$1.414^2 \leq 2 \leq 1.415^2,$$

a statement that can be tested without already "knowing" the value of $\sqrt{2}$.

When we want to deal with real numbers (and in particular, with irrational numbers), we almost always end up dealing with inequalities and approximations rather than actual equations. Thus, we are going to spend some time reviewing how exactly inequalities may be manipulated.

A word on things to come: the "skyscraper" of analysis is all about inequalities. However, once we get to the "cloud castle" of calculus, we will be back to caring mostly about equations. Thus, somehow, in the process of climbing to the top of the skyscraper, the inequalities get translated back into equalities. This is done using rules like the following:

Theorem. (to be proved later in the course) Let x be a real number. If we want to show that $x = 0$, it suffices to show the following: for every positive number ε,

$$|x| < \varepsilon.$$

Typically, when you see the symbol ε (Greek letter epsilon), you should think "small positive number." This is purely psychological: the statement would be just as correct if you replaced every ε with a y. Nevertheless, this "psychological" choice of variable can provide an important guide for intuition. When you see a statement like the theorem above, you should get the following idea:

> "If we can do a good enough job of showing that x is really close to zero, we'll know that x is actually equal to zero."

2 Rules for manipulating inequalities

If you read Section 0.2 of the textbook, you'll see a lot of talk about "solving" inequalities. The homework problems will use this term, so you'll need to make sure you understand what the authors mean by it. However, I prefer to think of "manipulating" inequalities rather than "solving" them. For instance, if you use the authors' methods to "solve" the inequality

$$x^2 < 2,$$

you'll get something like

$$-\sqrt{2} < x < \sqrt{2}.$$

However, since $\sqrt{2}$ is hard to calculate, the initial inequality may be easier to work with than the "solved" version.

Nevertheless, the basic tools are the same whether you want to "solve" inequalities or simply "manipulate" them. Unfortunately, these tools, i.e., rules for manipulation, tend to be more about nitpicking than interesting ideas. I've distributed a handout of rules that you should use for reference. Here are a few "traps" you may be tempted to run into, if you're used to solving equations rather than inequalities:

- "I can multiply both sides by the same number." **ISSUE:** You need to check the sign first. If you're multiplying by a positive number, you're fine. But if you're multiplying by a negative number, you need to reverse the direction of the inequality sign.

- "I can square both sides." **ISSUE:** This only works if both sides are positive.

- "If I have an inequality like $(x-a)(x-b) < 0$, where a product is compared to zero, I can split it into the factors: $x-a < 0$, $x-b < 0$." **ISSUE:** What you can actually say in this particular case is that $x-a$ and $x-b$ have opposite signs. In other words, one is positive and the other is negative. Quadratic inequalities are more complicated than quadratic equations.

If there's an "interesting idea" in manipulating inequalities, it's this: in some situations (for instance, in quadratic inequalities), we divide into cases, connected by words like AND and OR. At this point, we are not only doing algebraic manipulations. We are also playing around with the *logical* relationships among the different inequalities. Here's an example:

Example. (Example 3, Section 0.2 in text) Consider the inequality $x^2 - x < 6$. Much as in the case of quadratic equations, we start out by making one side zero and factoring the other side:

$$\begin{aligned} x^2 - x &< 6 \\ x^2 - x - 6 &< 0 \qquad \text{(subtract 6 from both sides)} \\ (x-3)(x+2) &< 0 \qquad \text{(factor)} \end{aligned}$$

Now, this single inequality is equivalent to the statement that $x - 3$ and $x + 2$ have opposite signs. In other words,

$$\begin{gathered} \big((x-3<0) \text{ AND } (x+2>0)\big) \text{ OR} \\ \big((x-3>0) \text{ AND } (x+2<0)\big) \end{gathered}$$

We analyze these two cases separately.

Case 1:

$$\begin{aligned} x - 3 &< 0 \qquad \text{AND} \qquad & x + 2 &> 0 \\ x &< 3 & x &> -2. \end{aligned}$$

A shorthand for ($x < 3$ AND $x > -2$) is

$$-2 < x < 3.$$

Case 2:

$$\begin{aligned} x - 3 &> 0 \qquad \text{AND} \qquad & x + 2 &< 0 \\ x &> 3 & x &< -2. \end{aligned}$$

There are no values of x such that $x > 3$ AND $x < -2$. A shorthand for a statement that is never true is $0 = 1$.

Combining the two cases: We see that our initial inequality is equivalent to the statement

$$-2 < x < 3 \quad \text{OR} \quad 0 = 1.$$

In other words, it is true precisely when at least one of the following is true:

(i) $-2 < x < 3$

(ii) $0 = 1$

Since $0 = 1$ is *never* true, it follows that $x^2 - x < 6$ precisely when $-2 < x < 3$.

3 Assignment due Friday, September 30

Read "A Bit of Logic" and "Quantifiers" on pp. 4–6.
Problem Set 0.1, numbers 45, 46, 63, and 64. Problems 45 and 46 will be graded carefully.

Read pp. 8–9.
Problem Set 0.2, numbers 3, 4, and 12. Problems 4 and 12 will be graded carefully. DO NOT use the quadratic formula on problem 12.

Bonus Exercise. *Show that the following three conditions on x are equivalent:*

(i) $x < \sqrt{2}$.

(ii) $x^2 < 2$ OR $x < 0$.

(iii) There exists y such that ($x < y$ AND $y^2 < 2$).

Math 131, Lecture 3

Charles Staats

30 September 2011

1 Statements and conditions

A mathematical *statement* is either true or false. For instance, $0 < 1$ is true, while $0 = 1$ is false.

We can build statements out of other statements using the logical operators AND, OR, and NOT.

Statement	In Words	True or False?
$(0 < 1)$ OR $(0 = 1)$	At least one of the two statements $(0 < 1)$, $(0 = 1)$ is true.	True
$(0 < 1)$ AND $(0 = 1)$	Both of the statements $(0 < 1)$, $(0 = 1)$ are true.	False
NOT $(0 < 1)$	The statement $(0 < 1)$ is false.	False

Sometimes a statement may involve a variable. The statement $x < 1$ is either true or false, but we can't tell which until someone tells us what x is. This sort of statement might be called a *condition* on x.

Two conditions on x are *equivalent* if they hold for exactly the same values of x. For instance, the condition $x \neq 0$ is equivalent to the condition $\big((x > 0)$ OR $(x < 0)\big)$, since in both cases, the statement is true precisely when x is nonzero. There are several ways to say this:

$P(x)$ is equivalent to $Q(x)$.
$P(x)$ if and only if $Q(x)$.
$P(x) \iff Q(x)$

They all mean the same thing.

2 Functions: writing f instead of $\sqrt{}$

[Note: the following history is partly fictional. But it could have happened this way, and in my opinion, it's a lot more interesting to think about it like this than just to go through a dry "definition of a function."]

For many centuries, algebra was, essentially, the study of formulas. Periodically, when dealing with formulas, people would pose a problem that could not be solved using existing formulas. For instance, to the ancient Greeks, such a problem was, "What is the side length of a square with area 2?" They knew that the answer would be a solution to the equation $x^2 = 2$. Unfortunately, this presented a dilemma, since they had no formulas to solve such an equation.

There were, roughly speaking, two approaches to this dilemma. One approach, which was taken by Diophantus of Alexandria in the third century A.D., was to accept that certain equations have no solutions, and then try to determine which equations had solutions and which did not. Diophantus produced some marvelous mathematics this way, and the sorts of questions he asked have become important in many areas—for instance, in modern cryptography.

Unfortunately, Diophantus' marvelous mathematics was little comfort to the farmer who wanted to know how long his fence should be to get a square corral with a given area. The other approach, which might have been more useful to said farmer, was to say, "well, since we don't have a formula for this, let's invent one—and then figure out how to calculate it." Thus, the square root was born.

As the centuries progressed, mathematicians continued to add new notation to their formulas—exponential, logarithm, sine and cosine, and others. But eventually, this approach stopped working. In studying differential equations, the variety of solutions became so great that it was wholly impractical to invent a new notation for every type of solution. Thus, they started using the same notation, $f(x)$, for many different "formulas." They might say something like, "Let f be defined as the solution to the differential equation under consideration," and then proceed to use f as though it were $\sqrt{}$. Later on, they might use the same letter f for the solution to a different equation.

In mathematics, notation is usually just notation. But sometimes, a new notation can lead to new insights. For instance, the symbol 0 was originally introduced as a placeholder, so that one could write down numbers like 101. But once the symbol was introduced, people began to realize that it made sense to think of zero as a number—a conceptual breakthrough.

In the case at hand, mathematicians began to realize that they could study the "set of all things that can be written as f." In trying to understand what these "things that can be written as f" really were, they came up with the following definition.

Definition. A *function* f is a rule that, given a number x, outputs a number $f(x)$.

Let's consider the case of the square root function f, defined by $f(x) = \sqrt{x}$ (or, if you prefer, defined by $f = \sqrt{}$). We would like to define f as follows:

> For each number x, the function f assigns to x that number y such that $y^2 = x$.

Unfortunately, this definition has a couple problems:

- This definition is ambiguous. For instance, if $x = 1$, then $f(x)$ could be either 1 or -1. To resolve this ambiguity, we require that $f(x)$ be nonnegative.

- If x is negative, there is *no* number y such that $y^2 = x$; in this case, $f(x)$ is undefined.

To resolve these difficulties, we make the following, better definition:

> For each *nonnegative* number x, the function f assigns to x the unique *nonnegative* number y such that $y^2 = x$.

The second difficulty, in particular, illustrates an important fact: a function may be defined on only *some* real numbers.

Definition. The *domain* of a function is the set of all numbers x such that $f(x)$ is defined.

Definition. Let f and g be functions. We say that $f = g$ if f and g have the same domain, and for every value of x in that domain, $f(x) = g(x)$.

Warning. *If f is a function, it may be tempting to write something like*

$$f = x^2 + 1.$$

This "equation" makes no sense. f is a function, whereas $x^2 + 1$ is a number (even if we're not sure which number it is). It does not make any sense to ask whether a function is equal to a number; they are simply different kinds of objects. If you use this sort of sloppy notation on homework or tests, you will *lose points for it.*

3 Composing functions

In the functions above, I always used x for a variable. There is nothing special about x; the square root function can be defined by $f(t) = \sqrt{t}$ just as easily as $f(x) = \sqrt{x}$. More importantly, we can plug in other things for a variable—numbers, other variables, expressions, even other functions. For instance, if f is defined by $f(x) = x^2$, then we may write things like

$$\begin{aligned} f(-2) &= (-2)^2 = 4 \\ f(x+t) &= (x+t)^2 = x^2 + 2tx + t^2 \\ f(x^2) &= (x^2)^2 = x^4. \end{aligned}$$

Note that none of these is a *definition* for f; they are all *consequences* of the definition that $f(x) = x^2$.

If f and g are both functions, then we may define a new function, denoted $f \circ g$, by

$$(f \circ g)(x) = f(g(x)).$$

This is called the *composition* of f and g; it is read "f composed with g."

Example. Let f be the function $x \mapsto x^2$, and let g be the function $x \mapsto x^2 + 1$. Compute $f \circ f$, $f \circ g$, and $g \circ f$.

Solution. $f \circ f$ is defined by

$$(f \circ f)(x) = f(x^2) = (x^2)^2 = x^4.$$

$f \circ g$ is defined by

$$(f \circ g)(x) = f(x^2 + 1) = (x^2 + 1)^2 = x^4 + 2x^2 + 1.$$

$g \circ f$ is defined by

$$(g \circ f)(x) = g(x^2) = (x^2)^2 + 1 = x^4 + 1.$$

We could just as well have computed $g \circ f$ by

$$(g \circ f)(x) = (f(x))^2 + 1 = (x^2)^2 + 1 = x^4 + 1. \qquad \square$$

Note that functional composition is not commutative: in the example above, $f \circ g$ is not equal to $g \circ f$.

4 Piecewise-defined functions

The easiest way to define a function is, of course, using an algebraic formula. But in a way, this misses the whole point of functions—that they can be used for rules that *cannot* be described using an algebraic formula. For instance, it is a (not entirely obvious) fact that the "rule"

$$f(x) = \text{the number } y \text{ such that } y^5 + 20y + 16 = x \tag{1}$$

gives an unambiguous definition for a function f, and a (much less obvious) fact that this function f cannot be expressed in terms of simpler algebraic functions like r^{th} roots. Even though we can't "solve" the equation $y^5 + 20y + 16 = x$ for y in terms of x in the sense of giving a formula, one *can* show that the solution exists as a function.

Another way to define functions is using a combination of logic and algebra. For instance, the following are two perfectly good functions:

$$f(x) = \begin{cases} x^3 - 7 & \text{if } x < 1, \\ x^2 & \text{if } x \geq 1. \end{cases}$$

$$g(x) = \begin{cases} -1 & \text{if } x \text{ is rational,} \\ 1 & \text{if } x \text{ is irrational.} \end{cases}$$

You may have a tendency to think that functions like these are somehow less "real" than functions defined using only algebra. You need to get past this. Piecewise-defined functions, as they are called, are just as "real" as functions given by algebraic formulae, and are extremely important. For instance, if you want to actually compute approximate values for the function defined in (1) above, your best bet may well be to construct a piecewise-defined function that is very close to it, and compute the values of that function.

An extremely important piecewise-defined function is the absolute value function.

Definition. The *absolute value function* is defined by

$$f(x) = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

$f(x)$ is typically denoted $|x|$, read "the absolute value of x." Note that $|x|$ is always nonnegative.

Example. Let f be the function defined by

$$f(x) = \begin{cases} x+1 & \text{if } x \leq 1 \\ x-1 & \text{if } x > 1. \end{cases}$$

"Solve" the inequality

$$f(2x) \leq x + 1.$$

Solution. Since f is piecewise-defined, we split into cases.

Case 1: $2x \leq 1$. In this case, the inequality reads $2x + 1 \leq x + 1$, so the case is equivalent to

$$\begin{array}{ccc} 2x \leq 1 & \text{AND} & 2x+1 \leq x+1 \\ x \leq \frac{1}{2} & \text{AND} & x \leq 0 \\ & x \leq 0. & \end{array}$$

Case 2: $2x > 1$. In this case, the inequality reads $2x - 1 \leq x + 1$, so the case is equivalent to

$$\begin{array}{ccc} 2x > 1 & \text{AND} & 2x-1 \leq x+1 \\ x > \frac{1}{2} & \text{AND} & x-1 \leq 1 \\ x > \frac{1}{2} & \text{AND} & x \leq 2 \\ & \frac{1}{2} < x \leq 2. & \end{array}$$

Combining the two cases, we see that the condition $f(2x) \leq x + 1$ is equivalent to the condition

$$x \leq 0 \text{ OR } \tfrac{1}{2} < x \leq 2. \qquad \square$$

5 Digression: Completing the square (avoiding the quadratic formula)

The quadratic formula is, in my opinion, drastically overemphasized in most algebra courses. It is rather ridiculous that people who have not studied math in thirty years might walk around remembering some (probably wrong) variant of "minus b plus or minus the square root of b squared minus four ac all over two a" without any recollection of why this is significant. Thus, I am going to *forbid* you to use the quadratic formula on anything you turn in (including tests and homework). Instead, I will expect you to use the technique of *completing the square*, which is a much more powerful idea that is in fact used to *derive* the quadratic formula. It's also easier to remember, in that the only formula involved is $(b/2)^2$.

Example. (Example 13 in the book.) "Solve" the inequality $x^2 - 2x - 4 < 0$. DO NOT use the quadratic formula.

Solution. Recall the important process of *completing the square:* to complete the square of $x^2 \pm bx$, add $(\frac{1}{2}b)^2$. In our case, $b = -2$, so $(\frac{1}{2}b)^2 = (-1)^2 = 1$. So, we need to turn the left side into $x^2 - 2x + 1$. We do this by adding 5 to both sides.

$$\begin{aligned} x^2 - 2x - 4 &< 0 \\ x^2 - 2x + 1 &< 5 \\ (x-1)^2 &< 5 \\ |x - 1| &< \sqrt{|5|} = \sqrt{5} \end{aligned}$$

Thus,

$$\begin{aligned} -\sqrt{5} < {} & x - 1 < \sqrt{5} \\ 1 - \sqrt{5} < {} & \quad x \quad < 1 + \sqrt{5}. \end{aligned}$$

□

6 Assignment: Due Monday, October 3

Section 0.2, problems 45 and 46. DO NOT use the quadratic formula, contrary to the book's instructions. Problem 46 will be graded carefully.

Skim Section 0.3 (pp. 16–22). Do the Concepts Review on p. 22 (answers on p. 24) to see if you need to read the section more closely; don't hand this in.
You may want to look at Example 3, p. 18. This process of completing the square is important. Make sure you understand it.
Do Section 0.3, problems 17, 18, 23, and 24. Problems 18 and 24 will be graded carefully.

Section 0.5, problem 2. This problem will be graded carefully.

Math 131, Lecture 4

Charles Staats

3 October 2011

1 Working with absolute values

Recall the absolute value function,

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

When asked to "solve" an inequality or an equation involving absolute values, it is *always* possible to get rid of the absolute values by splitting into cases. However, this can be ridiculously involved. The first pair of absolute value signs gives us two cases. If there is a second pair of absolute value signs, then each of these two cases splits into two subcases, for a total of four subcases. If there is a third occurrence of an absolute value, we end up with eight subsubcases. And so on.

For this reason, we often try to take "shortcuts," using rules for manipulating absolute values.

The rules for multiplication and division are easy:

$$|ab| = |a||b|$$
$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$$

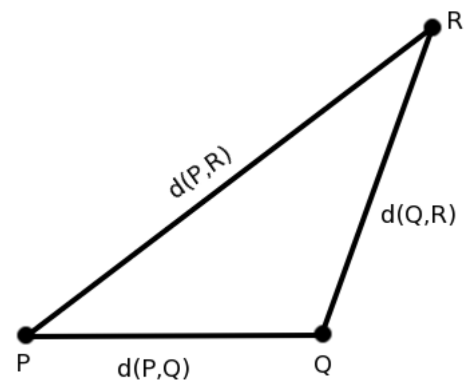
If we want to do addition or subtraction, the rules are not nearly so nice. We end up with inequalities rather than equations:

$$|a+b| \leq |a| + |b|$$
$$|a-b| \geq |a| - |b|.$$

The addition rule can be used to deduce the subtraction rule: If we set $a = c$ and $b = d - c$, the subtraction rule gives

$$\begin{aligned} |a+b| &\leq |a| + |b| \\ |c + (d-c)| &\leq |c| + |d-c| \\ |d| &\leq |c| + |d-c| \\ |d| - |c| &\leq |d-c|. \\ |d-c| &\geq |d| - |c|. \end{aligned}$$

The textbook calls the addition rule the "Triangle Inequality." This term is properly reserved for another inequality. Consider three points $P, Q,$ and R. Let $d(P, Q)$ denote the distance from P to Q.



The standard fact that "the shortest distance between any two points is a line" tells us that

$$d(P, R) \leq d(P, Q) + d(Q, R).$$

If $a, b,$ and c are real numbers, they also represent points on the number line. Moreover, the distance between a and b is precisely $|b - a|$, and so the triangle inequality for absolute values is

$$|c - a| \leq |b - a| + |c - b|.$$

We can deduce the addition rule from this: Let $a = 0$, $b = \alpha$, and $c = \alpha + \beta$. These substitutions were chosen precisely so that

$$\begin{aligned} c - a &= \alpha + \beta \\ b - a &= \alpha \\ c - b &= \beta. \end{aligned}$$

Thus, the triangle inequality gives us

$$|\alpha + \beta| \leq |\alpha| + |\beta|,$$

which is the addition rule.

2 Graphing functions

One of the keystones of modern mathematics is the interaction between algebra and geometry via the graphing of equations. In some cases, one can use algebra to prove a geometric result; you may have seen this sort of analysis used in analyzing the conic sections. However, in this course, we will be going mostly in the opposite direction: we will be using the geometry to gain additional insight about the algebra. See, for instance, the discussion of the triangle inequality above.

The basic approach to graphing functions is, of course, quite simple:

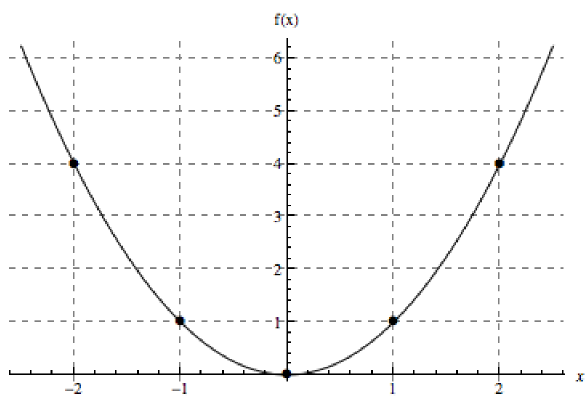
1. Choose some values of x.
2. Calculate and plot the points $(x, f(x))$.
3. "Connect the dots."

Example. Graph the function f defined by $f(x) = x^2$.

Solution. We first calculate f at a few points:

x	f(x)
-2	4
-1	1
0	0
1	1
2	4

Now, we plot these points and "connect the dots":



And, in this case, it works like a charm! □

Question: How do I know when I've plotted enough points?

You don't—not really. Later on, we'll discuss how to show definitively that you've plotted enough points, but no one ever does this in real life. But here are some general guidelines. They're not guaranteed to work, but they usually do if you're smart about it.

1. *Make sure it is "clear" how to connect the dots.* If your points are too far apart, either vertically or horizontally, you may need to plot some more. Generally speaking, you want the graph to be going "up" or "down" for several points at a time.

2. *Use your knowledge of the function.* If the graph you've drawn is a line, then the function had better be equal to a function of the form $f(x) = mx + b$; if it's not, then you probably need to plot some more points.

   If your function has an $(x - a)$ in the denominator, then you are dividing by zero at $x = a$. So, you probably want to plot extra points near $x = a$.

3. *Test some extra points in between the one you've already plotted.* When you think you know what the graph looks like, plot a few more points in between the ones you've already plotted. If they are about where your drawing says they should be, that's a good sign.

Question: How do I know I've got all the interesting features of the graph?

The best answer to this is to use calculus. Since we can't do that yet, it may be helpful to try to figure out what the function looks like in the "boring" part. Most of the functions we will give explicit formulas for this quarter will look like ax^n for very positive and very negative values of x. When the function starts looking like this, there's a good chance you're in the "boring" part.

You do probably want to make sure you get all the x-intercepts, i.e., all the points where $f(x) = 0$.

Issue: Discontinuities; undefined points

You probably want to figure out what the function's "natural domain" is, i.e., where it is defined. Make sure to figure out what is going on at the "edges" of this natural domain. If the domain can be written in interval notation, see what's going on near the (non-infinite) endpoints of all the intervals

If the function is piecewise-defined, you usually *don't* want to try to "connect the dots" between different pieces.

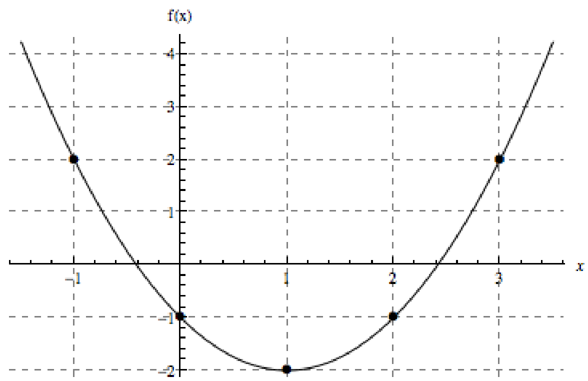
Things that can go right

For the most part, the discussions above focus on things that can go wrong. Sometimes there are also things that can be helpful. For instance, lines are very easy (more on this in a bit).

Other important techniques include translations. If you can write $f(x)$ as $g(x) + c$, where c is a constant, then the graph of $f(x)$ can be obtained from the graph of $g(x)$ by translating up by c. This can be useful, because g might be nicer algebraically than f. If you can write $f(x) = g(x - c)$, then the graph of f is obtained from the graph of g by translating g to the right by c.

Example. Graph the function f defined by $f(x) = (x - 1)^2 - 2$.

Solution. If $g(x) = x^2$, then $f(x) = g(x - 1) - 2$. Thus, take the graph of g, and translate it one to the right and down two.



$\square$

Example. Recall the inequality from the end of the last lecture:

$$f(2x) \leq x + 1,$$

where

$$f(x) = \begin{cases} x+1 & \text{if } x \leq 1 \\ x-1 & \text{if } x > 1. \end{cases}$$

Graph the functions $g(x) = f(2x)$ and $h(x) = x + 1$. Use the resulting graph to study the set of values of x satisfying the inequality $g(x) \leq h(x)$.

3 Assignment 3 due Wednesday, October 5

Section 0.2, problems 53, 54, 57, and 63. Problems 54 and 63 will be graded carefully.

Section 0.5, problem 13. This will be graded carefully.

Section 0.6, problems 13-16. Problems 14 and 16 will be graded carefully.

Math 131, Lecture 5

Charles Staats

5 October 2011

1 Quantifiers; rescuing a mess-up from last lecture

Last lecture, I introduced the following two inequalities:

$$|a+b| \le |a| + |b| \qquad \text{“Addition Rule”}$$
$$|a-b| \ge |a| - |b|. \qquad \text{“Subtraction Rule”}$$

I then did an somewhat abysmal job of explaining how the Addition Rule implies the Subtraction Rule. I'm going to see if I can do any better the second time around. I'll also try to use this as an excuse to discuss quantifiers. Please let me know immediately if I start talking gibberish again.

Example. Take the following statement as given: For every pair of real numbers a and b,

$$|a+b| \le |a| + |b|.$$

Use it to prove the Subtraction Rule: For every pair of real numbers a and b,

$$|a-b| \ge |a| - |b|.$$

Solution. The Addition Rule applies to every pair of real numbers. We've chosen to write this pair as a and b. But if a and b are a pair of real numbers, so are b and $a-b$. Applying the Addition Rule to the pair $b, a-b$, we obtain

$$\begin{aligned} |b+(a-b)| &\le |b| + |a-b| \\ |a| &\le |b| + |a-b| \\ |a| - |b| &\le |a-b| \\ |a-b| &\ge |a| - |b|. \end{aligned}$$

□

Let's review the logic here. The Subtraction Rule has the appearance of a condition on a and b: that $|a+b| \ge |a| - |b|$. Let's call this condition $P(a,b)$. Like all conditions, $P(a,b)$ is either true or false, but in principle, we don't know which until someone tells us what a and b are.

However, we want to show that this condition $P(a,b)$ holds for *every* possible choice of a and b. Statements of the form for all x, the condition $P(x)$ holds

will be increasingly common as we progress into the study of limits, continuity, and ultimately derivatives. The part of the statement "for all x" is called a *quantifier*. It may seem more reasonable to talk about "the quantifier" when we write the statement in symbols:

$$\forall x, P(x),$$

where $\forall$ stands for "for all." In this case, $\forall$ is the quantifier. The other important quantifier is $\exists$, which stands for "there exists." It showed up in one of the quiz questions yesterday:

> Let x be a positive real number. Show that **there exists** another positive real number y such that $y < x$.

When you are asked to prove a statement involving quantifiers, there's a typical narrative structure that is involved. It's easier to describe for the $\forall$ quantifier. If you are asked to prove that

> for every positive real number x, $P(x)$,

the proof typically starts out something like this:

> Let x be a positive real number. We'll show that $P(x)$ is true.

An important note here is that when you say, "Let x be a . . . ," *you don't get to choose x*. If it helps, imagine that someone else—an "opponent" or "enemy"—is going to try to find an x to spite you. What you are doing for the rest of the proof is showing that, no matter what x they choose, $P(x)$ holds.

The narrative structure for a $\exists$ proof is a bit more confusing, because the way you *tell* the proof is usually in the opposite order from the way you *figure out* the proof. If you're going to prove that

$$\exists y \text{ such that } y \text{ is irrational},$$

the proof you *tell* is probably going to have two steps:

1. Here's a specific number y that I've dreamed up. For instance, $y = \sqrt{2}$.
2. Here's why this specific y is irrational.

The trouble is, when you are *figuring out* the proof, it is often not clear what y you should pick. You have to wrestle with the condition on y until you have some y that you know (or at least suspect) works. And all of this initial work gets left out of the story you tell.

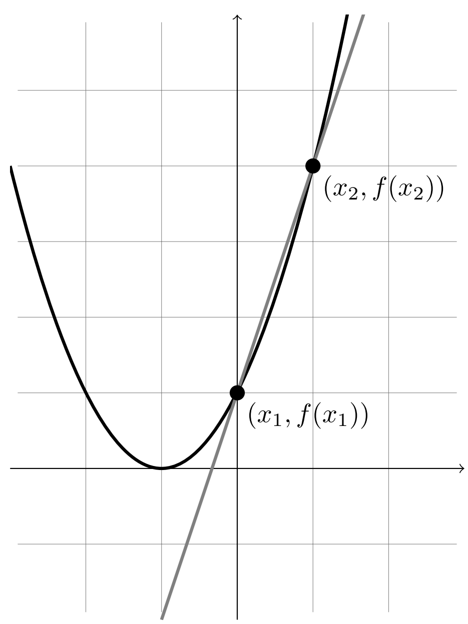
2 Walking on clouds: What does this formula mean?

At this point, I'm supposed to start motivating the notion of a "limit," as we ease toward the dreaded ε-δ definition. The problem is that, with the functions we know how to use right now, there aren't any really interesting or unexpected limits—they tend to be a bit boring. Certainly, there is no real need for this elaborate and potentially confusing definition. One of the few interesting examples is the difference quotient, not because it is hard to evaluate, but because without understanding limits well, it is easy to be skeptical that the answer has any meaning. One philosopher once called it the "ghost of a departed quantity."

Suppose we are considering the function

$$f(x) = (x+1)^2.$$

The graph of f is as follows:



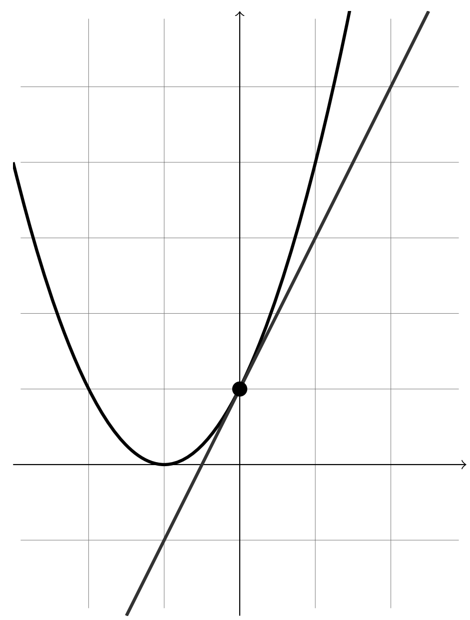
Given any two *distinct* values x_1, x_2 for x, there is a unique line through the two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. We may calculate its slope as follows:

$$\begin{aligned}
\text{slope} &= \frac{\text{change in } y}{\text{change in } x} \\
&= \frac{f(x_2) - f(x_2)}{x_2 - x_1} \\
&= \frac{(x_2+1)^2 - (x_1+1)^2}{x_2 - x_1}.
\end{aligned}$$

Let's see what happens when $x_1 = 0$, so the first point is $(0, f(0)) = (0, 1)$:

$$\begin{aligned} \text{slope} &= \frac{(x_2+1)^2 - 1}{x_2 - 0} \\ &= \frac{x_2^2 + 2x_2 + 1 - 1}{x_2} \\ &= \frac{x_2^2 + 2x_2}{x_2} \\ &= \frac{x_2(x_2+2)}{x_2} \\ &= x_2 + 2 \end{aligned}$$

The interesting thing is that this formula can be evaluated when $x_2 = 0$, even though the original formula cannot. Moreover, if you look at the graph



you can get an idea that this is probably the slope of the *tangent* line at $(0, 1)$. This is something of a miracle, since the original reasoning for the formula breaks down completely:

1. In the original formula, we're dividing by zero. How can this possibly give a meaningful answer?

2. If we look at how we derived the formula, it's supposed to give us the "slope of the line through $(x_1, f(x_1))$ and $(x_2, f(x_2))$." But if $(x_2, f(x_2))$ is the same point as $(x_1, f(x_1))$, this is ambiguous: there are infinitely many lines through these "two" points, all having different slopes. So where does this choice of a single line with slope 2 come from?

The original inventors of calculus more or less just accepted these miracles. They used this formula to "walk on clouds," even though they could not really explain why dividing by zero could give a reasonable answer—or any answer at all. And while they did great things this way, there were always mathematicians who distrusted these "miracles." Even more significantly, once people began looking at arbitrary functions, they realized that these "miracles" don't always work; and it became necessary to understand this miracle so they could know when it would work.

The solution to this turned out to be the definition of the *limit*, and was one of the most important steps in building the "skyscraper" to support the "cloud castle" of calculus.

3 Assignment 4 (due Friday, 7 October)

Section 0.2, problems 35, 36, 37, 38, 39, and 40. Problems 38 and 40 will be graded carefully.

Solve each of the following inequalities two different ways:

(a) By factoring and then dividing into cases.

(b) By completing the sqare.

Make sure you get the same answer both ways.

1. $x^2 - 1 \leq 0$
2. $x^2 - 4x + 3 < 0$
3. $x^2 + 2x - 3 \geq 0$
4. $x^2 + 2x - 3 > 0$
5. $x^2 - x - 6 \leq 0$
6. $x^2 - 3x - 28 \geq 0$

Problems 2 and 6 will be graded carefully.

Bonus Exercise. *(Not due until Wednesday, 12 October; worth three points) Find a piecewise-constant function f, defined on all x for which $-2 < x < 2$, such that for all such x,*

$$|f(x) - x^2| < 1.$$

Essentially, you are trying to approximate the squaring function by an easier-to-calculate function.

Remember—you need to actually give a proof that your function f works. This will involve showing, on each "piece" on which $f(x)$ is a constant c, that $|c - x^2| < 1$ on that "piece." (You'll never be able to find one c that works for all x, which is why you need several "pieces.")

Math 131, Lecture 6

Charles Staats

7 October 2011

1 Contrapositives in "real life"

Mathematics is often described as the "language of science." But the truth is, mathematics has its own language, based on informal[1] logic. It is vital to be fluent in this language in order to really understand mathematics. But this language is not useful only in mathematics. A good understanding of the language of informal logic is useful any time you want to debate anyone about anything.

In mathematics, one of the crucial skills is to be able to move facilely among different ways of saying the same thing. Very often, something that seems odd, or at least non-obvious, when stated one way, will become obviously true (or perhaps obviously false) when you move to a different way of saying them. Other times, you can take a true statement, and understand it better by thinking about different ways of saying it.

Right now, we will be concentrating on if-then statements. (These are sometimes called *conditionals*, but I'm going to avoid this term[2] because it is so similar to the term "condition" that I used for "conditions on x.")

Consider the statement

> Whatever does not kill you, makes you stronger.

At first glance, this does not look like an if-then statement. But we can make it into one:

> If something does not kill you, then it makes you stronger.

One advantage of an if-then statement is that we can take the contrapositive, which is formally equivalent. Thus, the statement above is equivalent to

> If something does not make you stronger, then it kills you.

Or, translating back into the original language,

[1] Roughly speaking, "informal logic" uses words and sentences and is comprehensible to humans. "Formal logic" uses symbols and is comprehensible to computers. The two-column proofs you may have done in high school geometry fall somewhere in between.

[2] Except for this one remark, of course.

> Whatever does not make you stronger, kills you.

Here we've taken a statement that seemed reasonable, and showed that it means the same thing as another statement that may seem absurd. This could be a neat trick if you are in a debate with someone who is defending the first statement.

Now, suppose we're trying to defend the first statement. In order to do that, we have to think about the second statement in a way that makes it seem more plausible. For one thing, we replace the too-general notion of "Everything" by the more restricted notion of "every trial," which is basically what we meant anyway. Then, we do something like the following.

> Every trial either kills you or makes you stronger. Thus, the only way a trial can fail to make you stronger is by killing you.

One additional note: From a careful, logical perspective, the statement "Whatever trial does not kill you, makes you stronger" does not exclude the possibility that some trial might *both* kill you *and* make you stronger. We typically fill in this bit of information from common sense, and there is nothing wrong with that in everyday language. But in mathematics, it is extremely important to be able *identify* when we are filling in information from "common sense" rather than logic, for two reasons. First, someone else's "common sense" might differ from ours, in which case we need to be able to defend our claims with logic. Second, "common sense" is sometimes very wrong, as we will see later in the course.

2 Saddling the Infinity beast

There is a great beast called infinity. This beast roams the plains of thought, going places that most people can only wonder at. It has always been one of the great mysteries. One of the greatest privileges of being a mathematician is to be able to sit astride this beast, to go where it goes and see the wondrous spectacles it sees.

Recall the quiz question that amounted to

> Prove that there is no least positive real number.

Some people said something like the following:

> The positive real numbers can be divided infinitely small, so there is no smallest.

This is essentially the right idea, but the explanation is wrong. There are no "infinitely small" numbers, so this notion of "dividing infinitely small" does not make sense. Metaphorically, you're headed in the right direction, but you've spooked the Infinity beast, and it won't carry you.

To hold yourself up, you need a saddle—something between you and the Infinity beast that will keep you from falling off. One of the best saddles, and the one that works here, is the notion of the *arbitrary*. You'd like to deal with "infinitely small" numbers, but that does not work. Instead, you deal with *arbitrarily small* numbers. Rather than trying to do things with infinitely small numbers, you take a finite number, often called ε, and do something that works "for arbitrarily small ε." In other words, what you are doing will still work, no matter how small ε gets. This "arbitrarily small" number ε is your "saddle," the "cushion" between you and the Infinity beast that lets you ride in comfort.

An important note here, worth repeating, is that you don't get to choose ε. If you choose a particular ε, whether it be $\varepsilon = 1$ or $\varepsilon = .000001$, you are limiting how small ε can be. In order for ε to be *arbitrarily small*, and hence provide a connection to the infinitesimal[3], you have to remember that ε is just, well, ε. Generally speaking, it's a small positive real number. You can't pretend you know more about it than that.

Coming back to the quiz question, no matter how small you make ε, $\varepsilon/2$ is always smaller. Thus, no matter how small you make ε, there is always something smaller.

3 Order of quantifiers

Last lecture, I discussed a bit about the quantifiers "for all" and "there exists." Let's consider how these are used in the statement of one of the quiz problems:

> Suppose that x is a positive real number. Show that there is another positive real number y such that $y < x$.

I mentioned in the last lecture that this includes the quantifier "there exists." But if you think about it, there are actually two quantifiers. The statement could be re-written more symbolically as

$$\forall x > 0, \exists y > 0 \text{ s.t. } y < x,$$

where "s.t." stands for "such that." If you imagine playing a game against an opponent, the opponent gets to move first (he gets to choose x), and then you get to move, choosing y in response. As we've seen, no matter what x he chooses, you can always choose $y = \frac{1}{2}x$ and win.

Let's see what happens if we reverse the order of the quantifiers and let you move first:

$$\exists y > 0 \text{ s.t. } \forall x > 0, y < x.$$

In other words, "there exists a positive number less than all positive numbers." When you play this game, you lose: whatever y you choose, your opponent then chooses $x = y$. Since $x = y$, it is *not* true that $y < x$, which is what you would have needed in order to win.

This example illustrates something important about the order of quantifiers: it's easier to "win" if the $\exists$ quantifier comes last—in other words, if you move

[3] "Infinite" means infinitely large. "Infinitesimal" means infinitely small. A woman who confused these two once told a speaker, "I enjoyed your lecture very much and thought it was of absolutely infinitesimal value." This was not the compliment she apparently thought it was.

second, rather than first. If this seems somewhat counterintuitive, think about it as a debate: it's easier to win if you get the last word.

As a final point, let's consider one formal definition of a limit:

Definition. Let f be a function whose domain includes all $x > m$, for some finite m. We say that

$$\lim_{x \to \infty} f(x) = c$$

if

$$\forall \varepsilon > 0, \exists M > m \text{ such that if } x > M, \text{ then } |f(x) - c| < \varepsilon.$$

I don't expect you to really understand this definition yet; we'll go over it more carefully next lecture, with more motivation. The key point to realize is this: if you look at the quantifiers, you will notice that the $\exists$ quantifier comes second. This means that in "playing the game," *you get the last word.* This does not guarantee that you will win, but it does give you a fighting chance.

Assignment $4\frac{1}{2}$ (due Monday, 10 October)

"Problems" 1 and 3 on page 33 of the book you will find at `http://www.phy.duke.edu/~rgb/Class/intro_physics_1/intro_physics_1.pdf`. These "problems" involve doing some reading—three times—and writing a couple of short essays. This portion of the book gives advice on how to learn. It's by one of my favorite professors when I was a college student. (He was much better at giving out candy than I am.) The essays will be collected and graded (by the instructor).

One final note: The assignment will be, quite literally, impossible to complete unless you start it by Friday, since part of the assignment is to work on it on three different days.

Assignment 5 (due Wednesday, 12 October)

- Consider the following two problems from the quiz:
  - Suppose that x is a positive real number. Show that there is another positive real number y such that $y < x$.
  - Either state the smallest positive real number, or prove that there is no such thing.

  Explain why these are really the same problem. This problem will be graded carefully.

- For each of the following statements,
  (a) Rewrite it as an if-then statement.
  (b) Give the contrapositive. (You may use "bad" as an abbreviation for "not good.") Optionally, rewrite it in a form resembling the original, rather than the "if-then" form.

  Here are the statements:
  1. Everyone with a beer has an ID.
  2. Everyone on the plane has a ticket.
  3. Every good boy does fine.[4]
  4. Every good boy deserves fudge.
  5. Good men die young.
  6. Only good men die young. [Note: This one is tricky.]

  Statements 2, 4, and 6 will be graded carefully.

- A friend of yours does not understand the formal definition of a limit. Write a paragraph explaining it to him so that it makes sense. (You may want to try this at least twice—once before the lecture on Monday, and once afterwards. Ideally, you should turn in several versions that show how your own understanding has improved.)

Bonus Exercise. *(worth three points) Find a piecewise-constant function f, defined on all x for which $-2 < x < 2$, such that for all such x,*

$$|f(x) - x^2| < 1.$$

Essentially, you are trying to approximate the squaring function by an easier-to-calculate function.

Remember—you need to actually give a proof that your function f works. This will involve showing, on each "piece" on which $f(x)$ is a constant c, that $|c - x^2| < 1$ on that "piece." (You'll never be able to find one c that works for all x, which is why you need several "pieces.")

[4] This sentence is used as a memory aide in music theory for the sequence of letters EGBDF, which partly explains why it is grammatically questionable.

Math 131, Lecture 7

Charles Staats

Monday, 10 October 2011

1 A bit of computer science

Suppose we have two computer programs for doing the same thing—say, multiplying matrices. You don't need to know what this means, except that matrices have size n. The bigger n is, the bigger the matrices are, and the longer it takes to multiply them.

Let's say the first computer program takes $100n^3$ operations to multiply two matrices of size n, while the second computer program takes n^4 operations to multiply the same two matrices. If we want to multiply matrices of size $n = 30$, let's see how long it takes:

$$\begin{array}{lrcccc} \text{First program:} & 100n^3 & = & 100 \cdot 30^3 & = & 2{,}700{,}000 \\ \text{Second program:} & n^4 & = & 30^4 & = & 810{,}000 \end{array}$$

Clearly, the first program is slower, taking 2.7 million operations instead of 810 thousand. On the other hand, speaking very loosely, a one gigahertz processor can execute a billion operations per second.[1] So, even the first program will only take .0027 seconds. The second program is faster, but the first is so fast already that no one will notice the difference.

On the other hand, suppose we want to multiply matrices of size $n = 500$:

$$\begin{array}{lrcccc} \text{First program:} & 100n^3 & = & 100 \cdot 500^3 & = & 12{,}500{,}000{,}000 \\ \text{Second program:} & n^4 & = & 500^4 & = & 62{,}500{,}000{,}000 \end{array}$$

In this case, on our one gigahertz processor, the first program takes 12.5 seconds, while the second takes 62.5 seconds. Not only is the first program faster in this case, but the times are long enough that we actually care. This illustrates a general point about computer science:

- When a program takes $f(n)$ operations to process something of size n, we generally only care what happens when n is big. When n is small, the program runs so quickly anyway that we don't care.[2]

[1] Please don't tell any computer science professor I said this. I'm speaking *very* loosely.
[2] This is a rule of thumb. There are plenty of exceptions.

Now, suppose you're a computer scientist, and someone hands you two programs and asks you which is faster? You see that the first program takes $f(n)$ operations, and the second program takes $g(n)$ operations. So, for any given value of n, you can simply calculate $f(n)$ and $g(n)$ and see which is smaller (hence faster).

The problem is, the person has told you absolutely nothing about what values of n they care about. Lacking this knowledge, you assume they probably care most about what happens when n is very large. Thus, you might consider the following criterion for telling them the f-program is faster than the g-program:

- For all large n, $f(n) < g(n)$. (1)

Unfortunately, it's not clear what you mean by "large n." One simple attempt would be to say something like "n is large if $n > 100$." This sort of idea would allow you to produce more precise versions of (1):

(i) For all $n > 100$, $f(n) < g(n)$.

(ii) For all $n > 1{,}000$, $f(n) < g(n)$.

(iii) For all $n > 10{,}000$, $f(n) < g(n)$.

Question. Which of these criteria is the hardest to prove / least likely to be true? Which implies the others?

If you notice, each of these versions of (1) has the following form: you first fix some N, and then take "n is large" to mean that "$n > N$." You then get a version of (1) that reads

- For all $n > N$, $f(n) < g(n)$.

Unfortunately, you have no idea how fast the person's computer is, so you have no idea what N to choose. So why not let it be arbitrary?

- There exists N such that for all $n > N$, $f(n) < g(n)$. (2)

If you recall our discussion of quantifiers as a "game," you may notice that we have just given ourselves an extra "move" by inserting a $\exists$ quantifier. Note that you've given yourself the first move rather than the "last word." You had no choice: the opponent's move of choosing an $n > N$ does not even make sense until you've told him what N is, so you have to go first.

Question. What happens if you give the opponent *both* moves?

2 Another saddle: the "sufficiently large"

If you think about it, we might have started out by asking, "What happens if $n = \infty$?" Unfortunately, asking whether $f(\infty) < g(\infty)$ will spook the Infinity beast so badly that we're likely to get gored. Instead, in (2), we've just stumbled upon another saddle that we can use to sit on the beast without actually touching it: the "sufficiently large." When a mathematician writes something like

For all sufficiently large n, $P(n)$ holds,

she means that we get to choose what we mean by "n is large." In symbols, the statement above would be written

$$\exists N \text{ s.t. } \forall n > N, P(n).$$

3 How many times faster?

Let's revisit computer science for a bit. Suppose we're comparing two computer programs for which the time required for an input of size n is, respectively,

$$f(n) = n + 1$$
$$g(n) = 2n.$$

It will turn out that for all n sufficiently large, $f(n) < g(n)$.

Question. How exactly do we show this?

Thus, by our previous notion, the f-program is faster. However, suppose we are asked the question, "How many times faster?" Let's consider the following table of values:

n	10	100	1,000	10,000
$\frac{f(n)}{g(n)}$	0.55	0.505	0.5005	0.50005

When $n = 10$, we see that the f-program takes 0.55 times as long as the g-program. By the time we get to $n = 10{,}000$, the f-program takes 0.50005 times as long as the g-program. It seems that the bigger n gets, the closer $f(n)/g(n)$ gets to one half. So, we might be tempted to say that "the value of the function f/g at ∞ is 0.5," and hence that in some sense, the f-program is twice as fast as the g program.

Unfortunately, if we say it this way, we've spooked the Infinity beast. So, let's try out our new "saddle" of the "sufficiently large":

- For all sufficiently large n, $\dfrac{f(n)}{g(n)} = \dfrac{1}{2}$.

Unfortunately, this is simply not true. No matter how big we make n, $f(n)/g(n)$ will never be exactly equal to 0.5. What seems to be true is something more like this:

- For all sufficiently large n, $f(n)/g(n)$ is close to $1/2$.

Again, the trouble is that we don't know what exactly we mean by "close." Perhaps we mean that $f(n)/g(n)$ is within .01, or .001, or .0001 of $1/2$. But it looks like it's even closer than this.

- For every possible notion of "close to $1/2$," for all sufficiently large n, $\dfrac{f(n)}{g(n)}$ is "close to $1/2$."

For an arbitrary (small) positive number ε, we get a notion of "close to $1/2$" as "within ε of $1/2$." Thus, in symbols, we may write the above as

$$\forall \varepsilon > 0,\ \exists N \text{ s.t. } \forall n > N,\ \left| \frac{f(n)}{g(n)} - \frac{1}{2} \right| < \varepsilon.$$

This leads us to the definition of a limit:

Definition. For a function h and a real number c, we say that c is the "limit of $h(n)$ as $n \to \infty$," written

$$c = \lim_{n\to\infty} h(n),$$

if

$$\forall \varepsilon > 0,\ \exists N \text{ s.t. } \forall n > N,\quad |h(n) - c| < \varepsilon.$$

This could also be rewritten as

$$\text{For all } \varepsilon > 0, \text{ there exists } N \text{ such that if } n > N, \text{ then } |h(n) - c| < \varepsilon.$$

Writing it this way makes my claim last lecture that "you get the last word" look a little bit better.

Assignment 5 (due Wednesday, 12 October)

- Consider the following two problems from the quiz:
  - Suppose that x is a positive real number. Show that there is another positive real number y such that $y < x$.
  - Either state the smallest positive real number, or prove that there is no such thing.

  Explain why these are really the same problem. This problem will be graded carefully.

- For each of the following statements,
  (a) Rewrite it as an if-then statement.
  (b) Give the contrapositive. (You may use "bad" as an abbreviation for "not good.") Optionally, rewrite it in a form resembling the original, rather than the "if-then" form.

  Here are the statements:
  1. Everyone with a beer has an ID.
  2. Everyone on the plane has a ticket.
  3. Every good boy does fine.[3]
  4. Every good boy deserves fudge.
  5. Good men die young.
  6. Only good men die young. [Note: This one is tricky.]

  Statements 2, 4, and 6 will be graded carefully.

- A friend of yours does not understand the formal definition of a limit. Write a paragraph explaining it to him so that it makes sense. (You may want to try this at least twice—once before the lecture on Monday, and once afterwards. Ideally, you should turn in several versions that show how your own understanding has improved.)

Bonus Exercise. *(worth three points) Find a piecewise-constant function f, defined on all x for which $-2 < x < 2$, such that for all such x,*

$$|f(x) - x^2| < 1.$$

Essentially, you are trying to approximate the squaring function by an easier-to-calculate function.

Remember—you need to actually give a proof that your function f works. This will involve showing, on each "piece" on which $f(x)$ is a constant c, that $|c - x^2| < 1$ on that "piece." (You'll never be able to find one c that works for all x, which is why you need several "pieces.")

[3] This sentence is used as a memory aide in music theory for the sequence of letters EGBDF, which partly explains why it is grammatically questionable.

Assignment 6 (due Friday, 14 October)

Section 1.5, problems 1-9. Remember: It is not enough to get the right answer. You have to convince the reader that your answer is right. Problems 2, 4, 6, and 8 will be graded carefully.

Math 131, Lecture 8

Charles Staats

Wednesday, 12 October 2011

1 Recalling the definition of a limit

First, let's recall the definition of a limit—both the informal and the formal versions.

We say that $\lim_{x\to\infty} f(x) = \ell$ if

Informal: For every version of "close to", we can choose some meaning for "large" such that if x is "large," then $f(x)$ is "close to" ℓ.

Formal: For all real $\varepsilon > 0$, there exists N such that for all $x > N$,

$$|f(x) - \ell| < \varepsilon.$$

The following table shows the correspondence between the informal version and the formal version.

Informal	Formal	Explanation
For every version of "close to"	For every $\varepsilon > 0$	Each ε gives us a meaning for "close to"—namely, "within ε."
we can choose some meaning for "large"	there exists N	When we've chosen N, we say that "large" means "bigger than N."
such that if x is "large,"	such that if $x > N$	As we've said, x is "large" if $x > N$.
then $f(x)$ is "close to" ℓ.	then $\|f(x) - \ell\| < \varepsilon$.	We've said "$f(x)$ is close to ℓ" should mean that "$f(x)$ is within ε of ℓ." Now, $\|f(x) - \ell\|$ is precisely the distance from $f(x)$ to ℓ, so saying "$f(x)$ is within ε of ℓ" is the same as saying that "$\|f(x) - \ell\| < \varepsilon$."

2 Newton's "definition of a limit"

Consider the following statement from Isaac Newton's seminal work, the *Philosophiae Naturalis Principia Mathematica*:

> Quantities, and the ratios of quantities, which in any finite time converge continually to equality, and before the end of that time approach nearer to each other than by any given difference, become ultimately equal.

This was centuries before mathematicians came up with the correct definition of a limit in order to build the "skyscraper" of analysis. Newton was trying to build his "cloud castle" of Calculus. It's kind of hard to see in the middle of a cloud, so it's no wonder he was confused: he thought he was proving a theorem rather than stating a definition.

Nevertheless, this statement has some of the key aspects of the definition of a limit. Newton understood that it is not enough just to say that one quantity "approaches" another. He put in a key phrase: approaches *nearer than by any given difference*. In other words, when we say that "$f(t)$ approaches ℓ," we really mean that $f(t)$ becomes *arbitrarily close* to ℓ. In more modern language, Newton's "difference" would probably be called ε. We would say that for any given ε, $f(t)$ must approach to within ε of ℓ.

And he incorporated another key understanding—how exactly does this "becoming close" depend on t? Newton saw t as *time*. What we called $f(t)$, he might have called "the value of f once the time t has passed." Letting t get larger is, for him, simply letting a lot of time pass. And when we think about it this way, we come to the following realization. In order for $f(t)$ to approach ℓ "nearer than a given difference," $f(t)$ must become nearer than that difference *in finite time*. In other words, there is a time N, after which $f(t)$ becomes—and remains—within ε of ℓ.

Thus, in Newton's language, we have the following definition of limit:

> We say that a function $f(t)$ (where t represents time) has a limit ℓ if for any given difference ε, within finite time, the quantity $f(t)$ approaches—and remains—nearer to ℓ than by ε.

Exercise. *Relate this definition to the formal definition of a limit by making a table like the one at the end of Section 1.*

3 Computing limits when they exist

One of the interesting things about limits (as well as other major characters we will meet in the study of Calculus) is that the usual methods of computing them look practically nothing like the definition. The following "theorem" (it's really a bunch of theorems stated at the same time) is essentially copied from page 68 of the textbook, and is quite useful for evaluating limits. It gives situations in which limits behave exactly as you might hope.

Theorem. ("Main Limit Theorem") In the following equations, if the right side makes sense, then the left side also makes sense and is equal to the right side.

1. $$\lim_{x\to\infty} k = k$$
2. $$\lim_{x\to\infty} \frac{1}{x} = 0$$
3. $$\lim_{x\to\infty} [f(x) + g(x)] = \left[\lim_{x\to\infty} f(x)\right] + \left[\lim_{x\to\infty} g(x)\right]$$
4. $$\lim_{x\to\infty} [f(x) - g(x)] = \left[\lim_{x\to\infty} f(x)\right] - \left[\lim_{x\to\infty} g(x)\right]$$
5. $$\lim_{x\to\infty} [f(x) \cdot g(x)] = \left[\lim_{x\to\infty} f(x)\right] \cdot \left[\lim_{x\to\infty} g(x)\right]$$
6. $$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to\infty} f(x)}{\lim\limits_{x\to\infty} g(x)}$$
7. $$\lim_{x\to\infty} [f(x)]^n = \left[\lim_{x\to\infty} f(x)\right]^n$$

"The right side makes sense" means, for now, that the limits in question exist (as real numbers) and there is no division by 0.

This theorem can be proved from the definition of the limit. The proofs are not even that difficult. But the only way they can ever be interesting is when you do them yourself. Watching someone else do them is terribly boring, so I'll skip the proofs—at least for now—and move straight to discussing how to use the theorem to actually compute limits.

Warning. *If you use this theorem (typically, repeated applications of this theorem) to compute a limit, then you will have shown, in the process, that the limit exists. However, if you try to apply this theorem, and end up with something that makes no sense, you will* not *have shown that the original limit does not exist.*

Example. (Example 2, p. 78 in the textbook) Compute

$$\lim_{x\to\infty} \frac{x}{1+x^2}.$$

In particular, show that it exists.

Solution. The most obvious thing to try here is to apply Rule 6, which would tell us that

$$\lim_{x\to\infty} \frac{x}{1+x^2} = \frac{\lim_{x\to\infty} x}{\lim_{x\to\infty} 1+x^2},$$

assuming that the righthand side makes sense. Unfortunately, the right hand side does not make sense: the limits on the righthand side do not exist.[1]

A more successful way to solve this problem is to first divide both the top and the bottom by the highest power of x that appears in the denominator.

$$\begin{aligned}
\lim_{x\to\infty} \frac{x}{1+x^2} &= \lim_{x\to\infty} \frac{x}{1+x^2} \cdot \frac{1/x^2}{1/x^2} && \text{(algebra)} \quad (1)\\
&= \lim_{x\to\infty} \frac{\frac{1}{x}}{\frac{1}{x^2}+1} && \text{(algebra)} \quad (2)\\
&= \frac{\lim\limits_{x\to\infty} \frac{1}{x}}{\lim\limits_{x\to\infty} [(\frac{1}{x})^2+1]} && \text{(Rule 6)} \quad (3)\\
&= \frac{\lim\limits_{x\to\infty} \frac{1}{x}}{\left(\lim\limits_{x\to\infty} \frac{1}{x}\right)^2 + \lim\limits_{x\to\infty} 1} && \text{(Rules 3, 7)} \quad (4)\\
&= \frac{0}{0^2+1} && \text{(Rules 2, 1)} \quad (5)\\
&= 0. && (6)
\end{aligned}$$

To the right of each line is written the justification: why do we know it is equal to the previous line (assuming it is defined)?

A few words should be said on how we actually know the limits exist. If we actually want to be careful here, our knowledge of the limits goes from the bottom of the stack of formulas to the top. Because line (5) makes sense, the theorem tells us that line (4) makes sense and is equal to it. Because line (4) makes sense, the theorem tells us that line (3) makes sense and is equal to it. And so on, all the way up to the top (which is what we cared about to begin with). □

General procedure for computing limits of rational functions:

A rational function, as you may recall, is a function of the form

$$f(x) = \frac{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0}{b_k x^k + b_{k-1}x^{k-1} + \cdots + b_1 x + b_0}.$$

When faced with a function like this and asked to compute $\lim_{x\to\infty} f(x)$, here is a procedure that often works:

1. Multiply the numerator and denominator both by $1/x^k$.
2. Use the rules of the "Main Limit Theorem" to "distribute" the limit signs. Bring them further and further "inside" the formula, until all the limits are of the form $\lim_{x\to\infty} 1/x = 0$ or $\lim_{x\to\infty} k = k$.

[1] In a more sophisticated point of view that we will adopt later, the numerator and the denominator are both ∞. But ∞/∞ still does not make sense, as we will discuss.

Assignment 6 (due Friday, 14 October)

Section 1.5, problems 1-9. Remember: It is not enough to get the right answer. You have to convince the reader that your answer is right. Problems 2, 4, 6, and 8 will be graded carefully.

Assignment 7 (due Monday, 17 October)

Do the exercise at the end of Lecture 8, Section 2 on Newton's "definition of a limit."

In the textbook, Section 1.5, Problems 15, 16, and 18. Problems 16 and 18 will be graded carefully.

Complete the attached worksheet on graphing piecewise-defined functions.

Math 131, Lecture 9

Charles Staats

Friday, 14 October 2011

1 Remembering the definition of the limit

[Note: In case you have not figured it out yet, the definition of the limit is *very* important, and it *will* be on the test.]

Recall the definition of the limit:

Definition. For a function f and a fixed number ℓ, we say that

$$\lim_{x\to\infty} f(x) = \ell$$

if

$$\forall\varepsilon > 0, \exists N \text{ s.t. if } x > N, \text{ then } |f(x) - \ell| < \varepsilon.$$

Less formally:

For arbitrarily small positive ε, for sufficiently large x, $f(x)$ is within ε of ℓ.

This definition involves using two different “saddles” for the Infinity beast: the saddle of the “arbitrarily small,” and the saddle of the “sufficiently large.” It is the interaction of these two “saddles” that makes the definition of the limit so intricate. In order to get the definition right, you have to get both of them, and they have to be in the right order.

It may help to think of the “quantifier game.” In this definition, there are *three* moves:

1. First, your opponent moves. He gets to choose ε. In other words, he gets to choose how small is “arbitrarily small.”

2. Then, you get to move, by choosing N. Informally, you get to say how large is “sufficiently large.” Your choice can depend on ε, which your opponent has already chosen. But it *cannot* depend on x, which has not been chosen yet.

3. Finally, the judge gets to go—to decide who wins. In a sense, he gets to choose x; then, he determines whether this x does what you said (in which case you win) or not.

To reiterate:

$$\text{For } \underbrace{\text{arbitrarily small positive } \varepsilon}_{\text{opponent's move}}, \text{ for } \underbrace{\text{sufficiently large}}_{\text{your move}} x, \underbrace{f(x) \text{ is within } \varepsilon \text{ of } \ell}_{\text{judge's decision}}.$$

More formally,

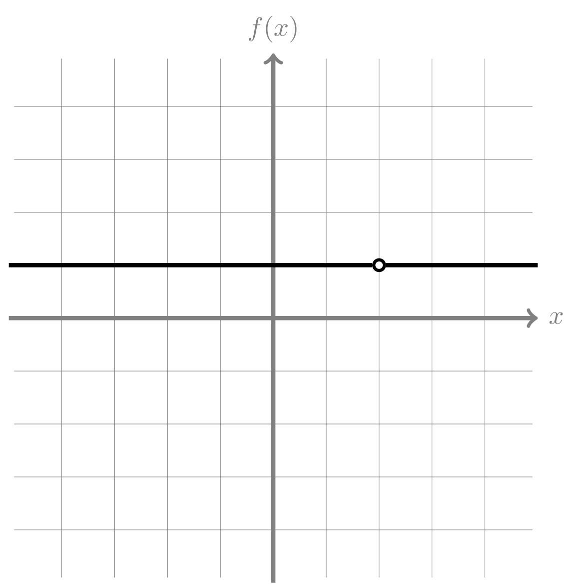
$$\underbrace{\forall \varepsilon > 0,}_{\substack{\text{opponent's}\\ \text{move}}} \quad \underbrace{\exists N}_{\substack{\text{your}\\ \text{move}}} \quad \text{s.t.} \quad \underbrace{\text{if } x > N, \text{ then } |f(x) - \ell| < \varepsilon.}_{\text{judge's decision}}$$

2 Limits as $x \to c$, where c is a real number rather than ∞

Consider the function f defined by

$$f(x) = \frac{x-2}{x-2}.$$

This function is not defined at $x = 2$, since computing it at that point would involve dividing by zero. However, for every value of x other than 2, $f(x) = 1$. Thus, the graph of f looks like this:



There is a "missing point" at $x = 2$, but it is clear what the value $f(2)$ ought to be. Mathematicians have a way of stating, precisely, that "the value of $f(x)$ at 2 ought to be 1": we write

$$\lim_{x\to 2} f(x) = 1.$$

More generally, for a given number c and a given value ℓ, we can claim that "as $x \to c$, $f(x) \to \ell$," which is also written

$$\lim_{x\to c} f(x) = \ell.$$

Note: When we write this notation, c is a value of x, whereas ℓ is a value of y. In particular, in our case,

$$\begin{aligned} c &= 2 \\ \ell &= 1. \end{aligned}$$

The definition of this sort of limit is quite similar to the definition we've already encountered. For the previous definition, we needed the concept of "arbitrarily large x" to deal let us talk about what happens at ∞ without spooking the Infinity beast. In a sense, this is the same thing as "arbitrarily close to ∞."

This time around, we're trying to understand what is (or should be) happening at c without actually touching c. (If you like, we want to avoid "spooking the c-beast.") The notion that replaces "arbitrarily large" is "arbitrarily close to c." Thus, we get the following definition.

Definition. We say

$$\lim_{x\to c} f(x) = \ell$$

if

For $\underbrace{\text{arbitrarily small } \varepsilon > 0}_{\text{opponent's move}}$, when x is $\underbrace{\text{sufficiently close}}_{\text{your move}}$ to c, $\underbrace{f(x) \text{ is within } \varepsilon \text{ of } \ell}_{\text{judge's decision}}$.

More formally,

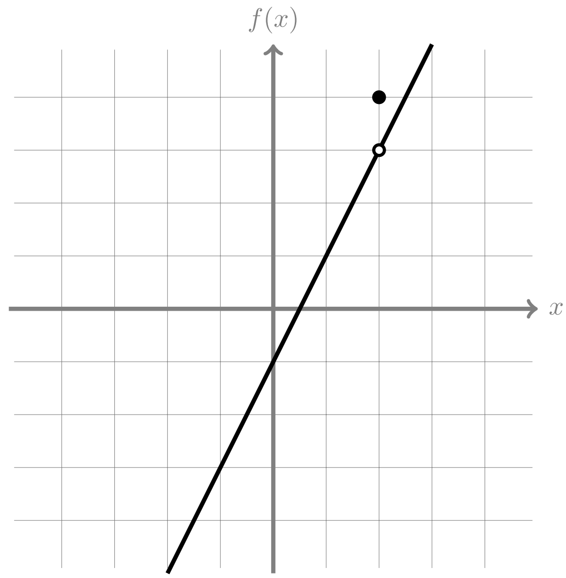
$$\underbrace{\forall \varepsilon > 0,}_{\substack{\text{opponent's}\\ \text{move}}} \quad \underbrace{\exists \delta > 0}_{\substack{\text{your}\\ \text{move}}} \text{ such that } \underbrace{\text{if } |x - c| < \delta \text{ and } x \neq c, \text{ then } |f(x) - \ell| < \varepsilon.}_{\text{judge's decision}}$$

The condition $x \neq c$ is added since we want to figure out what the value "should be" at c, without actually touching c.

Example. Consider the function f defined by

$$f(x) = \begin{cases} 2x - 1 & \text{if } x \neq 2, \\ 4 & \text{if } x = 2. \end{cases}$$

Its graph looks like this:



Let's use the ε-δ definition of the limit to show that

$$\lim_{x \to 2} f(x) = 3.$$

NOTE: When looking at this sort of example, we are not using the formal definition of the limit to better understand the function f. We are using the function f to better understand the definition. The formal definition becomes really useful when we are dealing with functions f for which we don't have formulas.

Solution. Let $\varepsilon > 0$ be given (by our opponent; we can't choose it). Before we go around choosing δ haphazardly to set what is "sufficiently close to 2," let's anticipate what the judge will say. In other words, let's "solve" the inequality he cares about as best we can, without knowing ε:

$$|f(x) - 3| < \varepsilon.$$

Since the judge does not care what happens when $x = c = 2$, we can assume

$f(x) = 2x - 1$. In this case, the judge's "test" is whether

$$\begin{aligned} |(2x-1)-3| &< \varepsilon \\ |2x-4| &< \varepsilon \\ 2|x-2| &< \varepsilon \\ |x-2| &< \tfrac{1}{2}\varepsilon. \end{aligned}$$

Now, we could proceed to finish "solving" the inequality; but in this case, that would be counterproductive. The condition we impose, by our choice of δ, is that

$$|x-2| < \delta.$$

Thus, if we set $\delta = \frac{1}{2}\varepsilon$, we are guaranteed that the judge will like all the "sufficiently small" values of x we allow him to look at. $\square$

Important: δ may depend on ε, and usually will. (Since our opponent has already chosen ε, we're allowed to use it.) But δ *cannot* depend on x. (The judge doesn't choose x until after we've already chosen δ.)

Assignment 7 (due Monday, 17 October)

Do the exercise at the end of Lecture 8, Section 2 on Newton's "definition of a limit."

In the textbook, Section 1.5, Problems 15, 16, and 18. Problems 16 and 18 will be graded carefully.

Complete the worksheet on graphing piecewise-defined functions.

Test Wednesday, 19 October

The test includes lectures through Wednesday, October 12, and assignments 1 through 7 (but not assignment 4.5). Note: The assignment numbers are one off from the lecture numbers, since I gave no (non-bonus) assignment on the first day of class. In particular, although the quiz does not include any new material from today's lecture, it does include the homework set due Monday.

Anything that appeared on a quiz will probably show up in some form on the test. (Exception: no contrapositive questions.) Anything that appeared on a non-bonus homework question might show up on the test.

If I said something in a lecture that did not make it in any form into a quiz or homework question, then it will not be on the test.

Math 131, Lecture 10

Charles Staats

Monday, 17 October 2011

1 Definition: Limits as $x \to c$

Recall from last time the definition of the limit of $f(x)$ as $x \to c$:

Definition. We say

$$\lim_{x \to c} f(x) = \ell$$

if

For $\underbrace{\text{arbitrarily small } \varepsilon > 0}_{\text{opponent's move}}$, when x is $\underbrace{\text{sufficiently close}}_{\text{our move}}$ to c, $\underbrace{f(x) \text{ is within } \varepsilon \text{ of } \ell}_{\text{judge's decision}}$.

More formally,

$$\underbrace{\forall \varepsilon > 0,}_{\substack{\text{opponent's} \\ \text{move}}} \quad \underbrace{\exists \delta > 0}_{\substack{\text{our} \\ \text{move}}} \text{ such that } \underbrace{\text{if } |x - c| < \delta \text{ and } x \neq c, \text{ then } |f(x) - \ell| < \varepsilon.}_{\text{judge's decision}}$$

A couple of notes on this definition:

- When we wanted to compute $\lim_{x \to \infty} f(x)$, this depended only on the function f. If we want to compute $\lim_{x \to c} f(x)$, we will, quite probably, have a different limit for every different choice of c.
- We specifically exclude the "judge" from looking at the value of $f(x)$ when $x = c$, because the limit is supposed to detect what "should be" going on at c without actually touching c. This was not an issue for $\lim_{x \to \infty} f(x)$, where we're sort of "setting $c = \infty$," because $f(\infty)$ does not make sense anyway.

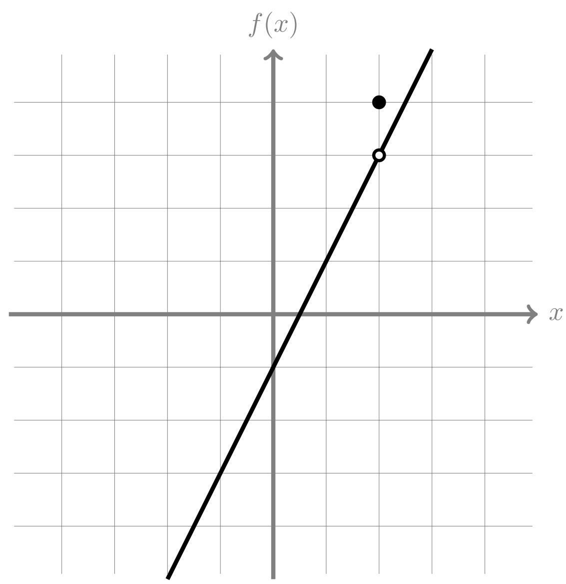
2 Examples: Using the ε-δ definition

At this point, we'll begin trying to understand the definition better by doing some examples of ε-δ proofs. We are doing this to help us understand the definition, and the concept, of limit, which is much more useful in more complicated situations.

Example. Consider the function f defined by

$$f(x) = \begin{cases} 2x - 1 & \text{if } x \neq 2, \\ 4 & \text{if } x = 2. \end{cases}$$

Its graph looks like this:



Let's use the ε-δ definition of the limit to show that

$$\lim_{x \to 2} f(x) = 3.$$

Note: When looking at this sort of example, we are not using the formal definition of the limit to better understand the function f. We are using the function f to better understand the definition. The formal definition becomes really useful when we are dealing with functions f for which we don't have formulas.

Let $\varepsilon > 0$ be given (by our opponent; we can't choose it). Before we go around choosing δ haphazardly to set what is "sufficiently close to 2," let's anticipate what the judge will say. In other words, let's "solve" the inequality he cares about as best we can, without knowing ε:

$$|f(x) - 3| < \varepsilon.$$

Since the judge does not care what happens when $x = c = 2$, we can assume $f(x) = 2x - 1$. In this case, the judge's "test" is whether

$$\begin{aligned}
|(2x-1)-3| &< \varepsilon \\
|2x-4| &< \varepsilon \\
2|x-2| &< \varepsilon \\
|x-2| &< \tfrac{1}{2}\varepsilon.
\end{aligned}$$

Now, we could proceed to finish "solving" the inequality; but in this case, that would be counterproductive. The condition we impose, by our choice of δ, is that

$$|x-2| < \delta.$$

Thus, if we set $\delta = \frac{1}{2}\varepsilon$, we are guaranteed that the judge will like all the "sufficiently small" values of x we allow him to look at.

> **Important:** δ may depend on ε, and usually will. (Since our opponent has already chosen ε, we're allowed to use it.) But δ *cannot* depend on x. (The judge doesn't choose x until after we've already chosen δ.)

If you recall the discussion of the "narrative" of a proof with quantifiers, the way you "tell" a proof is often quite different from the way you work it out. Now that we've worked out what the value of ε should be, let's "tell" the story of what goes on in the courtroom—without trying to get into the characters' heads (as we were, earlier, by anticipating the judge). Just the facts, ma'am.

Solution. Let $\varepsilon > 0$ be given. Set $\delta = \frac{1}{2}\varepsilon$. Assume $|x - 2| < \delta$ and $x \neq 2$.

Since $x \neq 2$, we know $f(x) = 2x - 1$. Hence,

$$\begin{aligned}
|f(x)-3| &= |2x-1-3| \\
&= |2x-4| \\
&= 2|x-2| \\
&< 2\delta \\
&= 2(\tfrac{1}{2}\varepsilon) \\
&= \varepsilon.
\end{aligned}$$

Thus, under these hypotheses, $|f(x) - 3| < \varepsilon$, as desired.

Hence, $\lim_{x\to 2} f(x) = 3$. □

Test Wednesday, 19 October

The test includes lectures through Wednesday, October 12, and assignments 1 through 7 (but not assignment 4.5). Note: The assignment numbers are one off from the lecture numbers, since I gave no (non-bonus) assignment on the first day of class. In particular, although the quiz does not include any new material from the Lectures 9 and 10, it does include the homework set due Monday, 17 October.

Anything that appeared on a quiz will probably show up in some form on the test. (Exception: no contrapositive questions.) Anything that appeared on a non-bonus homework question might show up on the test.

If I said something in a lecture that did not make it in any form into a quiz or homework question, then it will not be on the test.

Assignment 8 (due Friday, 21 October)

Give ε-δ proofs of the following facts:

$$\lim_{x\to 0} 7x = 0 \tag{1}$$

$$\lim_{x\to 1} 2x = 2 \tag{2}$$

$$\lim_{x\to -\frac{1}{2}} 4x + 1 = -1 \tag{3}$$

$$\lim_{x\to 5} \tfrac{1}{2}x - 2 = -\tfrac{1}{2} \tag{4}$$

They will all be graded carefully.

Math 131, Lecture 11

Charles Staats

Friday, 21 October 2011

1 Some notes on the test

First, some of you may be surprised, when you get back your test, to see how many points I may deduct even when you get the "correct answer." For instance, on the first problem, I gave a number of people scores of 4/10, even though they had the right "answer."

The way I see it is this: When I ask you a question, I'm not just asking you, "Where is London?" I'm asking you, "How do I get to London?" Someone who tells me to get to London by walking across the Atlantic Ocean, even though they have the right "answer" (London), will receive fewer points than someone who tells me to take a boat to Madrid. At least the second person will get me to the right continent without drowning.

Here are some common errors that were made on the test:

- The condition on x given by, for instance,
$$1 < x \quad \text{AND} \quad x < 2$$
can be abbreviated as
$$1 < x < 2.$$
The condition
$$1 < x \quad \text{OR} \quad x < 2$$
has *no such abbreviation*. These sorts of "combined statements" can *only* be used for AND. If you are dealing with an OR statement, you have to write it out in full, with the two statements included.

  [Incidentally, the particular OR statement above is actually equivalent to the statement $x = x$. Why?]

- $|2x + 7| > 5$ is an OR condition. $|2x + 7| < 5$ is an AND condition.

- When you complete the square on 2(b), you should end up with
$$(x + 2)^2 \geq 25.$$

At this point, since the right side is positive, you take the square root of both sides, getting

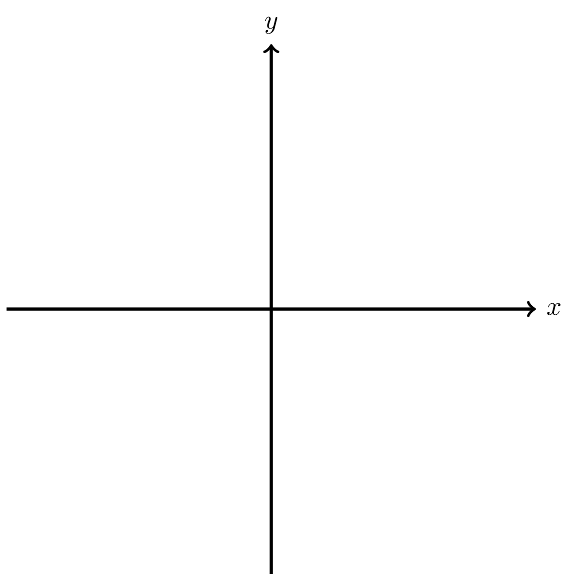
$$|x + 2| \geq 5.$$

You can then the rules for solving absolute value inequalities.

If you were solving an *equation*, you would probably want to use a $\pm$ sign rather than an absolute value. This does not work reliably for inequalities.

- It is certainly *possible* for $f(x)$ to equal its limit for large x. We just don't often look at such examples because they are not very *interesting*. However, when we study limits as $x \to c$, we *do* explicitly disregard what happens at $x = c$. It's important to distinguish between what happens to the y-values ($f(x)$ can equal c, although it does not have to) and the x-values (x cannot equal see—at least, as far as the judge is concerned).

- When we write either $\lim_{x\to\infty} f(x) = \ell$, the number ℓ is a value of y, not a value of x:



The same holds for the statement $\lim_{x\to c} f(x) = \ell$.

2 Limits as $x \to c$: an intuitive picture

It should come as a surprise to no one that a few minutes of the lecture today will be spent on the definition of the limit. Let's take a moment for a terribly imprecise, but intuitively useful, version:

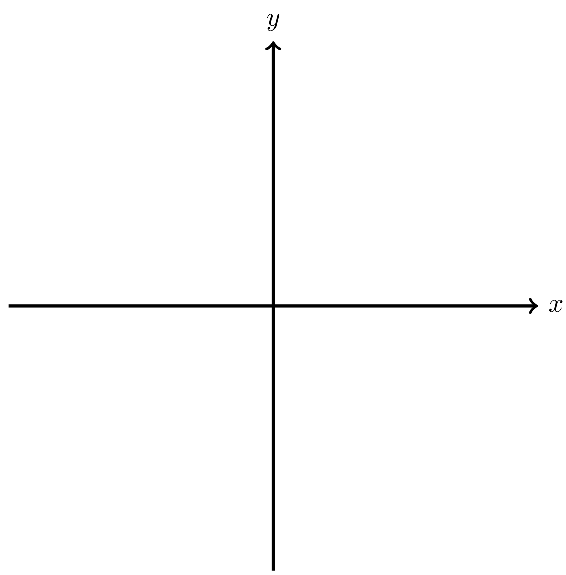
Definition. (Terribly Imprecise Version) We say that

$$\lim_{x \to c} f(x) = \ell$$

if the following holds:

> When x is close to c, then $f(x)$ is close to ℓ.

Here's the picture:



This relates to the more precise definitions as follows:

1. First, the opponent decides what it means to say "$f(x)$ is close to ℓ." He does this by choosing $\varepsilon > 0$, and then saying "$f(x)$ is close to ℓ" means precisely "$f(x)$ is within ε of ℓ."

2. Second, we, knowing what ε the opponent has chosen, get to decide what it means to say "x is close to c." We do this by choosing $\delta > 0$, and then saying "x is close to c" means precisely "x is within δ of c (but not equal to c)."

3. Finally, the judge takes our definitions and decides whether or not the basic statement is true: "When x is close to c, then $f(x)$ is close to ℓ." If it is true, we win; if not, the opponent wins.

Assignment 9 (due Monday, 24 October, 2011)

Read Section 1.1. If you aren't comfortable with sine and cosine, don't pay too much attention to the examples involving them. You do need to understand the picture in Example 6, however.

Section 1.1, Problems 29 and 30. These problems are (unusually) of the "answers only" variety. Problem 30 will be graded carefully.

Section 1.2, Problems 11–15. This time, the proof needs to be done carefully. Problems 12, 14, and 15 will be graded carefully.

Assignment 10 (due Wednesday, 26 October, 2011)

Consider the piecewise-linear function f defined by

$$f(x) = \begin{cases} 3x + 11 & \text{if } x \leq -3, \\ 1 - \frac{1}{3}x & \text{if } -3 < x \leq 3, \\ 3 - x & \text{if } x > 3. \end{cases}$$

1. Graph this function. Carefully.
2. Use the graph to "guess" what $\lim_{x \to -3} f(x)$ and $\lim_{x \to 3} f(x)$ are. If it looks like the two-sided limits don't exist, sleep on it, and then try graphing the function again.
3. Give ε-δ proofs that the two limits in the question above are what you say they are. Remember: You should have one proof for each (two-sided) limit. (Hint: you will probably want to let $\delta = \min\{\delta_1, \delta_2\}$, for appropriate values of δ_1 and δ_2.)

The third problem will be graded carefully.

Section 1.2, Problems 1–6. Problems 2, 4, and 6 will be graded carefully.

Section 1.3, Problems 1, 2, 5, and 6. Be sure to follow the instructions—these problems are about using the Main Limit Theorem carefully, not about finding the limits (those are very easy to guess). Problems 2 and 6 will be graded carefully.

Math 131, Lecture 12

Charles Staats

Monday, 24 October 2011

1 Ways limits can fail to exist

1.1 Jumps; one-sided limits

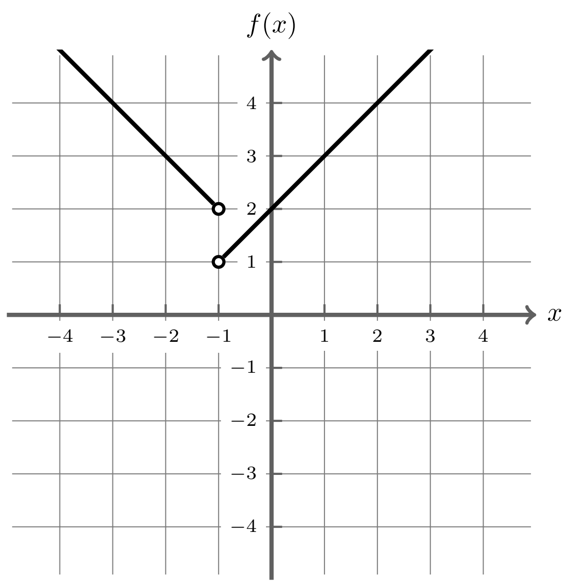
We consider the function f defined by

$$f(x) = \begin{cases} 1 - x & \text{if } x < -1, \\ 2 + x & \text{if } x > -1. \end{cases}$$

We have not defined this function at $x = -1$, but for the purpose of considering

$$\lim_{x \to -1} f(x),$$

f does not have to be defined at -1; and even if it is, we don't care what its value is.



In this situation, the limit does not exist. To handle "jumps" like this, we have the notion of one-sided limits.

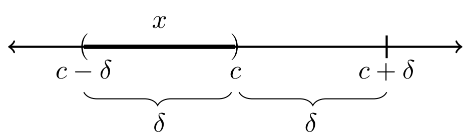
Definition. We say that "$f(x)$ approaches ℓ as x approaches c *from the left*," written

$$\lim_{x\to c^-} f(x) = \ell,$$

if

$$\forall\varepsilon > 0,\ \exists\delta > 0 \text{ s.t. if } \boxed{c-\delta < x < c,} \text{ then } |f(x)-\ell| < \varepsilon.$$

The boxed part says that "x is to the left of c and within δ of it":

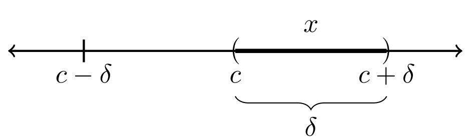


We say that "$f(x) \to \ell$ as $x \to c$ from the right," written

$$\lim_{x\to c^+} f(x) = \ell,$$

if

$$\forall\varepsilon > 0,\ \exists\delta > 0 \text{ s.t. if } \boxed{c < x < c+\delta}, \text{ then } |f(x)-\ell| < \varepsilon.$$



Exercise. *Using this ε-δ definition, show, for the function f defined above, that*

$$\lim_{x\to -1^-} f(x) = 2$$
$$\lim_{x\to -1^+} f(x) = 1.$$

Theorem. The two-sided limit $\lim_{x\to c} f(x)$ exists if and only if both the one-sided limits exist and are equal. In this case, we have

$$\lim_{x\to c^-} f(x) = \lim_{x\to c} f(x) = \lim_{x\to c^+} f(x).$$

This theorem is not *that* difficult to prove, but we will refrain because of time constraints. The basic idea is as follows: when the opponent gives us an ε, we

- Find a δ_1 that works for the left-hand limit.
- Find a δ_2 that works for the right-hand.
- Set $\delta = \min\{\delta_1, \delta_2\}$.

1.2 Infinite limits

We say $\lim_{x \to c} f(x) = \infty$ if

Informal: For arbitrarily $\boxed{\text{large } K,}$ when x is sufficiently close to c, then $\boxed{f(x) > K.}$

Formal: $\forall K,\ \exists \delta > 0$ s.t. if $0 < |x - c| < \delta,$ then $f(x) > K.$

We say $\lim_{x \to c} f(x) = -\infty$ if

Informal: For arbitarily $\boxed{\text{negative } K,}$ when x is sufficiently close to c, then $\boxed{f(x) < K.}$

Formal: $\forall K,\ \exists \delta > 0$ s.t. if $0 < |x - c| < \delta,$ then $f(x) < K.$

Example. If

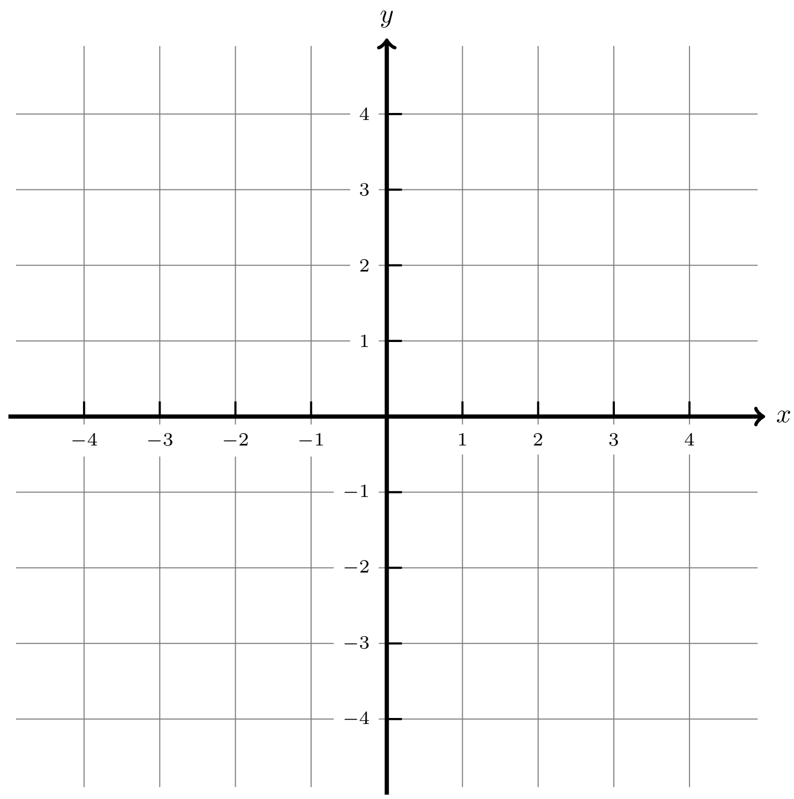
$$f(x) = \frac{1}{(x+1)^2} + \frac{1}{x-1},$$

then

$$\lim_{x \to -1} f(x) = \infty,$$

while $\lim_{x \to 1} f(x)$ does not exist in any sense. However,

$$\lim_{x \to 1^-} f(x) = -\infty$$
$$\lim_{x \to 1^+} f(x) = \infty.$$



1.3 Limits that just plain don't exist

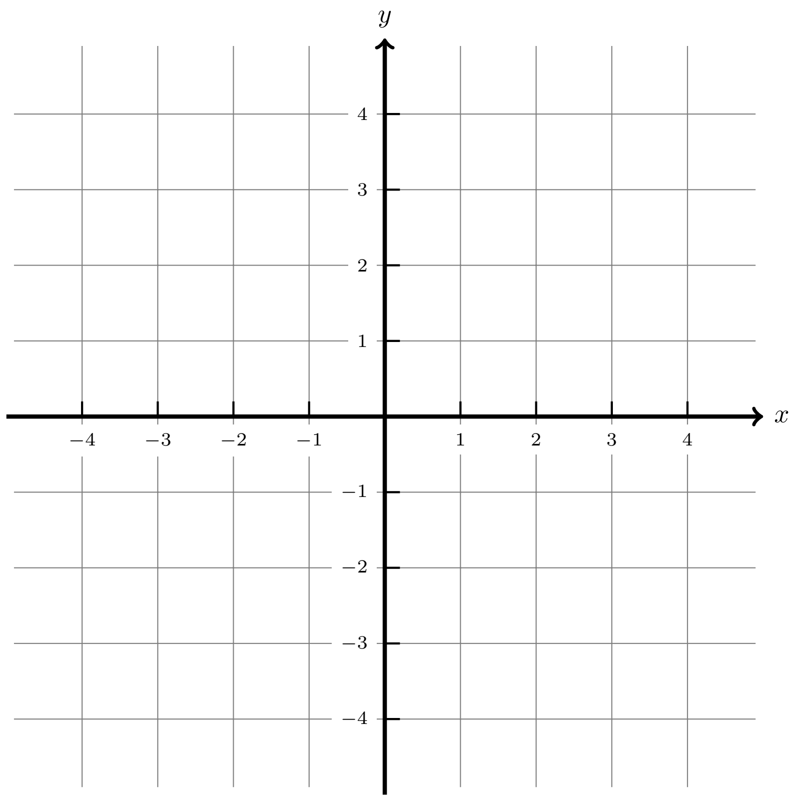
Consider the function g defined by

$$g(x) = \begin{cases} 0, & \text{if } x \text{ is rational,} \\ 1, & \text{if } x \text{ is irrational.} \end{cases}$$

Consider

$$\lim_{x \to 0} f(x).$$

No matter how small x is, there are always smaller values at which f is 0 (say, $x = \frac{1}{n}$, for some really big integer n) and smaller values at which f is 1 (say, $x = \frac{1}{n}\sqrt{2}$, for an even bigger integer n). So, no version of the limit as $x \to 0$ can exist—not the left-hand limit, not the right-hand limit, not even if we allow limits that are $\pm\infty$.



2 ε-δ proofs for more complicated functions

Example. Let f be the function defined by

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

Show that

$$\lim_{x\to 2} f(x) = 4.$$

First, we do a preliminary analysis. The definition of limit, applied in this specific instance, is as follows:

$$\forall\varepsilon > 0,\ \exists\delta > 0 \text{ such that if } 0 < |x-2| < \delta,\ \text{then } |x^2-4| < \varepsilon.$$

Let's look, specifically, at what the judge cares about:

$$\begin{aligned} |x^2-4| < \varepsilon \quad &\iff \quad |x-2||x+2| < \varepsilon \\ &\iff \quad |x-2| < \frac{\varepsilon}{|x+2|}. \end{aligned}$$

At this point, it might be tempting to say, "Let δ be $\varepsilon/|x+2|$." However, to set δ like this, we'd have to know x, and we don't: the judge does not choose x until after we've chosen δ.

This is the sort of thing that makes ε-δ proofs so much more difficult for nonlinear functions: We have to control x to make two things happen at once:

- Make sure $\varepsilon/|x+2|$ does not get too small.
- Make sure $|x-2|$ does not get too big.

The second task is easy—it's precisely what the δ is designed to do. But the first task is much harder, and it's where we have to start.

Suppose we know that $|x-2| < 1$. If you "solve" this inequality, you get precisely that $1 < x < 3$. Thus, $x+2$ is clearly positive, and so $|x+2| = x+2$. Thus, we have

$$\begin{aligned} 1 < x \quad & < 3 \\ 3 < x+2 \quad & < 5 \\ 3 < |x+2| \quad & < 5 \\ \frac{1}{3} > \frac{1}{|x+2|} & > \frac{1}{5} \\ \frac{\varepsilon}{3} > \frac{\varepsilon}{|x+2|} & > \frac{\varepsilon}{5}. \end{aligned}$$

Note: in the inequalities above, each line *implies* the next, but is *not* necessarily equivalent.

Thus, we have:

$$\text{If } |x-2| < 1, \text{ then } \frac{\varepsilon}{5} < \frac{\varepsilon}{|x+2|}.$$

As long as we choose $\delta \le 1$, we have some control over how small $\varepsilon/|x+2|$ can be. This takes care of the first, problematic task. Since we've done that, the second task is comparatively easy: we need to ensure that

$$|x-2| < \frac{\varepsilon}{5}.$$

This works as long as $\delta \le \varepsilon/5$.

So, to accomplish the first task, we need $\delta \le 1$. Once we've done this, to accomplish the second task, we need $\delta \le \varepsilon/5$. Since 1 and $\varepsilon/5$ are both positive numbers, we can simply set

$$\delta = \min\left\{1, \frac{\varepsilon}{5}\right\}.$$

Now, we've finally got a plan. Let's head into the courtroom and see what the judge says.

Proof. Let $\varepsilon > 0$ be given. Set

$$\delta = \min\left\{1, \frac{\varepsilon}{5}\right\}.$$

Now, suppose $0 < |x - 2| < \delta$. Then we have

$$|x^2 - 4| = |x - 2|x + 2.$$

To proceed further, we need to know something about $|x + 2|$. Since $\delta \leq 1$, we know

$$\begin{aligned} |x - 2| &< 1 \\ -1 < x - 2 &< 1 \\ 1 < x &< 3 \\ 3 < x + 2 &< 5, \end{aligned}$$

and consequently, $x + 2$ is positive.

$$3 < |x + 2| < 5.$$

Hence, $|x + 2|$ is a positive number less than 5. Thus, we have

$$\begin{aligned} |x^2 - 4| &= |x + 2||x - 2| \\ &< 5|x - 2| \\ &< 5\delta, \end{aligned}$$

since $|x - 2| < \delta$

$$< 5 \cdot \frac{\varepsilon}{5},$$

since $\delta \leq \varepsilon/5$

$$= \varepsilon.$$

If $|x - 2| < \delta$, then $|x^2 - 4| < \varepsilon$. Hence, the limit is as claimed. □

Bonus Exercise. *Show, using an ε-δ argument, that*

$$\lim_{x \to 3} x^2 = 9.$$

3 Main Limit Theorem

One important tool for computing limits is the "Main Limit Theorem." This is essentially the same as the one we discussed for limits as $x \to \infty$, but with the following addition:

$$\lim_{x \to c} x = c.$$

In practice, when we're computing limits as $x \to c$, repeated applications of the Main Limit Theorem usually end up just telling us that

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

When this holds for a function f, we say that f is *continuous at c.* (More on this next lecture.)

Assignment 10 (due Wednesday, 26 October, 2011)

Consider the piecewise-linear function f defined by

$$f(x) = \begin{cases} 3x + 11 & \text{if } x \le -3, \\ 1 - \frac{1}{3}x & \text{if } -3 < x \le 3, \\ 3 - x & \text{if } x > 3. \end{cases}$$

1. Graph this function. Carefully.
2. Use the graph to "guess" what $\lim_{x\to -3} f(x)$ and $\lim_{x\to 3} f(x)$ are. If it looks like the two-sided limits don't exist, sleep on it, and then try graphing the function again.
3. Give ε-δ proofs that the two limits in the question above are what you say they are. Remember: You should have one proof for each (two-sided) limit. (Hint: you will probably want to let $\delta = \min\{\delta_1, \delta_2\}$, for appropriate values of δ_1 and δ_2.)

The third problem will be graded carefully.

Section 1.2, Problems 1–6. Problems 2, 4, and 6 will be graded carefully.

Section 1.3, Problems 1, 2, 5, and 6. Be sure to follow the instructions—these problems are about using the Main Limit Theorem carefully, not about finding the limits (those are very easy to guess). Problems 2 and 6 will be graded carefully.

Assignment 11 (due Friday, 28 October, 2011)

- Section 1.5, Problems 51 and 52. Problem 52 will be graded carefully.
- Section 1.6, Problems 1, 3, and 5. None of these will be graded carefully.
- Do the exercise on page 2 of the notes for Lecture 12. This will be graded carefully.
- Let $a \ne 0$ and c be arbitrary real numbers. (In other words, the "opponent" gets to choose a and c, and he is allowed to choose anything as long as he does not set a equal to 0.) Give an ε-δ proof that

  $$\lim_{x\to c} ax = ac.$$

  This will be graded carefully.

- A "sequence" (a_n) is a list of numbers, for instance,

$$0, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \ldots.$$

Typically, the n^{th} term will be denoted a_n. Thus, in the sequence above, we have

$$\begin{aligned} a_1 &= 0 \\ a_2 &= \frac{1}{2} \\ a_3 &= \frac{3}{4} \\ a_4 &= \frac{7}{8} \\ a_5 &= \frac{15}{16} \\ &\vdots \\ a_n &= \frac{2^{n-1}-1}{2^{n-1}} \\ &\vdots \end{aligned}$$

Explain why a "sequence" is the same thing as a "function with domain the positive integers." [Once you've thought about this for long enough, it may become so obvious that you have very little to say. Unfortunately, the grader cannot give you credit just for writing "It's obvious."]

- Bonus: Do the Bonus Exercise on page 7 of the Lecture 12 notes.

Math 131, Lecture 13

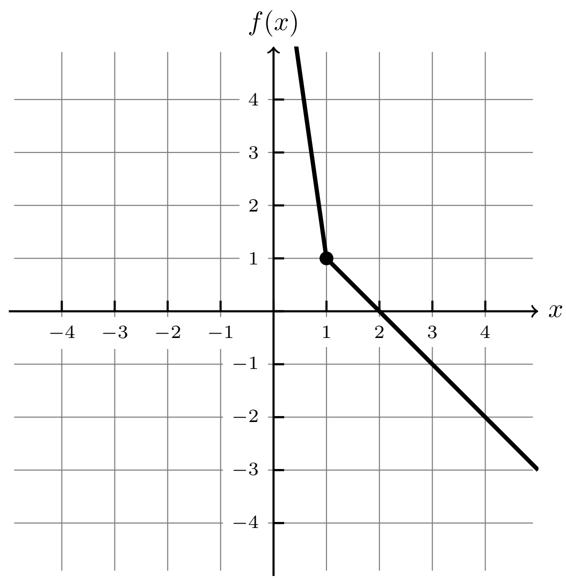
Charles Staats

Wednesday, 26 October 2011

1 An ε-δ proof for the limit of a piecewise-linear function

Example. Let f be the function defined by

$$f(x) = \begin{cases} -7x+8 & \text{if } x \leq 1, \\ 2-x & \text{if } x > 1. \end{cases}$$



Give an ε-δ proof that

$$\lim_{x \to 1} f(x) = 1.$$

Solution. First, since I don't feel like doing an extensive preliminary analysis, I'm going to let you in on a little trick: If you have a function that looks like

$$f(x) = mx + b,$$

with $m \neq 0$, and you want to do an ε-δ proof for f, then

$$\delta = \tfrac{1}{|m|}\varepsilon$$

is a good guess for δ. I'm not saying it will always work, but it's worth trying.

Now, for the function f in this particular example, we actually have two linear equations. To the left of $x = 1$, we have $f(x) = -7x + 8$. This suggests that for the left-hand limit, we should take

$$\delta_1 = \tfrac{1}{|-7|}\varepsilon = \tfrac{1}{7}\varepsilon.$$

To the right of $x = 1$, we have $f(x) = -x + 1$, which suggests that for the right-hand limit, we should take

$$\delta_2 = \tfrac{1}{|-1|}\varepsilon = \varepsilon.$$

Since we're interested in the two-sided limit, we should probably try

$$\delta = \min\{\delta_1, \delta_2\} = \min\{\tfrac{1}{7}\varepsilon, \varepsilon\} = \tfrac{1}{7}\varepsilon.$$

Now, let's proceed to the actual proof, and see if this choice of δ works out.

Proof. Let $\varepsilon > 0$ be given. Set $\delta = \frac{1}{7}\varepsilon$.

Assume $0 < |x - 1| < \delta$; we consider separately what happens when $x < 1$ and when $x > 1$.

Case 1: $x < 1$. Here, $f(x) = -7x + 8$, and so

$$\begin{aligned} |f(x) - \ell| &= |f(x) - 1| \qquad &= |-7x + 8 - 1| \\ &= |-7x + 7| \\ &= |7x - 7| \\ &= 7|x - 1| \\ &< 7\delta \\ &= 7 \cdot \tfrac{1}{7}\varepsilon \\ &= \varepsilon. \end{aligned}$$

Case 2: $x > 1$. Here, $f(x) = 2 - x$, and so

$$\begin{aligned} |f(x) - 1| &= |1 - x| \\ &= |x - 1| \\ &< \delta \\ &= \tfrac{1}{7}\varepsilon \\ &< \varepsilon. \end{aligned}$$

In either case, $|f(x) - \ell| = |f(x) - 1| < \varepsilon$, as desired.

$\square$

2 Continuity

Given what we've already seen, the simplest definition of continuity is the following:

Definition. A function f is said to be *continuous* at a point x_0 if

(i) f is defined at x_0, AND

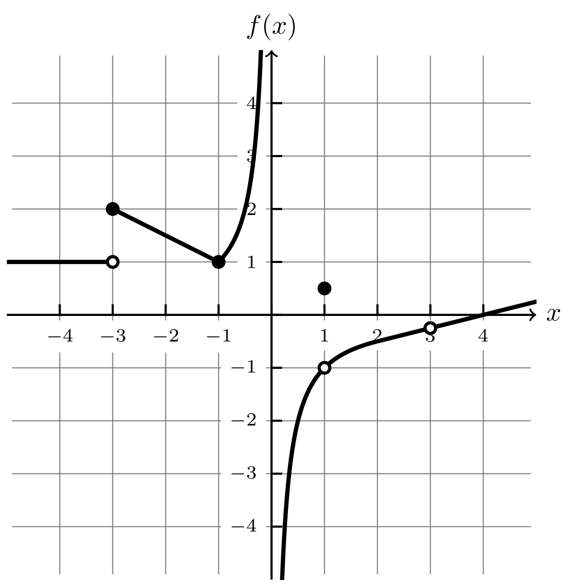
(ii) $\lim\limits_{x \to x_0} f(x)$ exists, AND

(iii) $\lim\limits_{x \to x_0} f(x) = f(x_0)$.

If f is continuous at every point of an interval, we say that f is *continuous on that interval.*

If f is continuous at every point in its domain, we may say simply that f is *continuous.*

Example. Consider the function f whose graph looks like this:



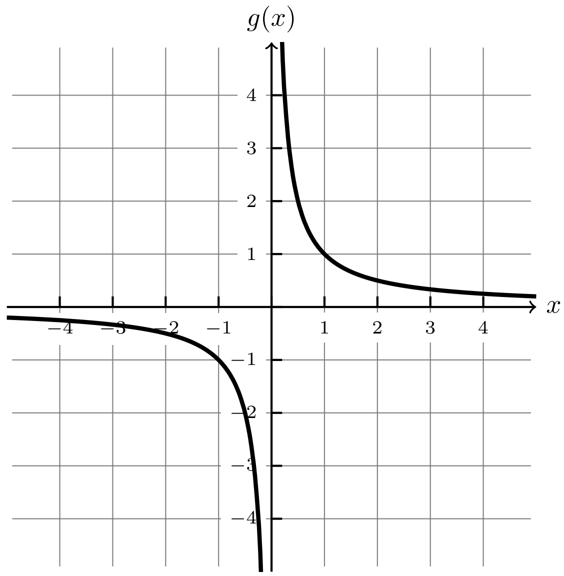
1. At which points on the closed interval $[-5, 5]$ is f not continuous?

f fails to be continuous at the x-values -3, 0, 1, and 3

2. At which points *in its domain* is f not continuous?

The x-values -3 and 1. The other x-values listed above do not lie in the domain of f.

Example. The function g defined by $g(x) = 1/x$

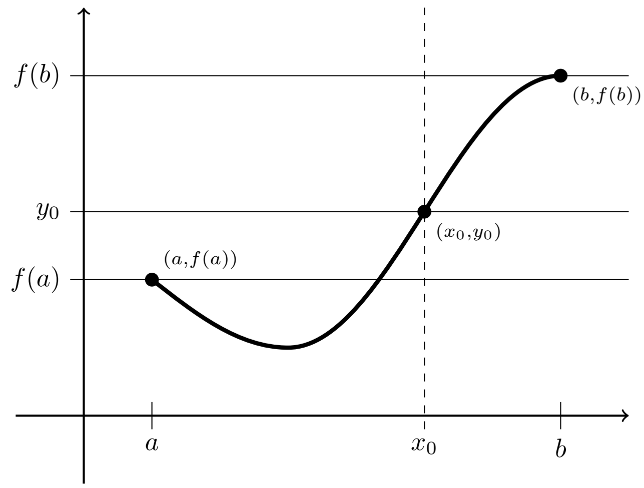


would be called continuous, since it is continuous *on its domain*; in other words, it is continuous on the intervals $(-\infty, 0)$ and $(0, \infty)$. [It is not continuous at $x = 0$, but this "does not count" because 0 is not a point of its domain; f is not defined at 0.]

The intuitive notion is that a function is continuous on an interval $[a, b]$ if you can draw f from the point $(a, f(a))$ to the point $(b, f(b))$ without picking up your pencil. This intuitive idea is, unfortunately, something of a dead end if we try to use this idea, rather than limits, to define what "continuous" ought to mean. However, the following theorem does seem to capture the notion that if you draw continuously from one point to another, you have to pass through all the points in between:

Theorem. (Intermediate Value Theorem) Suppose f is continuous on the closed interval $[a, b]$. Suppose we have a value y_0 such that $f(a) < y_0 < f(b)$. Then f hits the value y_0 somewhere on the open interval (a, b). In other words, there exists x_0 such that $a < x_0 < b$ and $f(x_0) = y_0$.

Here's the picture:



Intuitively, for the (continuous) f to go from the line $y = f(a)$ to the line $y = f(b)$, it has to pass through the line $y = y_0$ somewhere. That "somewhere" is our x_0.

We won't try to prove this right now.

Assignment 11 (due Friday, 28 October, 2011)

- Section 1.5, Problems 51 and 52. Problem 52 will be graded carefully.
- Section 1.6, Problems 1, 3, and 5. None of these will be graded carefully.
- Do the exercise on page 2 of the notes for Lecture 12. This will be graded carefully.
- Let $a \neq 0$ and c be real numbers (which you don't get to choose). Give an ε-δ proof that

  $$\lim_{x \to c} ax = ac.$$

  This will be graded carefully.
- A "sequence" (a_n) is a list of numbers, for instance,

  $$0, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \ldots .$$

  Typically, the n^{th} term will be denoted a_n. Thus, in the sequence above, we have

  $$\begin{aligned} a_1 &= 0 \\ a_2 &= \frac{1}{2} \\ a_3 &= \frac{3}{4} \\ a_4 &= \frac{7}{8} \\ a_5 &= \frac{15}{16} \\ &\vdots \\ a_n &= \frac{2^{n-1} - 1}{2^{n-1}} \\ &\vdots \end{aligned}$$

  Explain why a "sequence" is the same thing as a "function with domain the positive integers." [Once you've thought about this for long enough, it may become so obvious that you have very little to say. Unfortunately, the grader cannot give you credit just for writing "It's obvious."]
- Bonus: Do the Bonus Exercise on page 7 of the Lecture 12 notes. (This exercise comes at the end of Section 2 of the Lecture 12 notes. If you want to attempt it, you should probably read the entire section.)

Assignment 12 (due Monday, October 31, a.k.a. Halloween)

Section 1.3, Problems 3, 4, 7, and 8. Remember, these problems are about showing you understand how to use the Main Limit Theorem, not about finding the limits. Problems 4 and 8 will be graded carefully.

Section 1.6, Problems 2, 4, 6-8, 32, and 33. Problems 6-8 and 32 will be graded carefully. You do not need to give ε-δ proofs for these problems.

Bonus problem: Let f be a function (which you do not get to choose). Consider the statement

> For all x_0 in the domain of f, for all $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$.

Explain why this is equivalent to the statement that "f is continuous." (One dilemma you may need to address: Why does it make no difference if we write $|x - x_0| < \delta$ rather than $0 < |x - x_0| - \delta$?)

Math 131, Lecture 14

Charles Staats

Friday, 28 October 2011

1 Zeno's arrow paradox

Wikipedia has a nice summary of Zeno's arrow paradox:

> In the arrow paradox (also known as the fletcher's paradox), Zeno states that for motion to occur, an object must change the position which it occupies. He gives an example of an arrow in flight. He states that in any one (durationless) instant of time, the arrow is neither moving to where it is, nor to where it is not. It cannot move to where it is not, because no time elapses for it to move there; it cannot move to where it is, because it is already there. In other words, at every instant of time there is no motion occurring. If everything is motionless at every instant, and time is entirely composed of instants, then motion is impossible.

This more or less captures the central conceptual idea in differential calculus. When we have an object in motion, we'd like to be able to talk about how fast it is going at any given instant. But the essence of motion is moving from one position to another, whereas in a single, durationless instant, an object only occupies a single position. So how can we even think about the speed at a particular instant—or, to use slightly fancier terminology, the "instantaneous velocity"?

There are basically two ways to think about this. The more mathematically rigorous way is to use limits. The idea here is to say "since we can't make the change in time zero, let's make it arbitrarily small." Since we can't touch the (here) Zero beast, let's handle it through the saddle of the Arbitrary.

The other way—the "walking on clouds" approach that was used for the first two centuries or so after calculus was invented—is to say, in essence, "Let's pretend that the instant at time t_0 actually does have an 'infinitesimal' duration, which we call dt, and see what happens." This "infinitesimal" duration, dt, is bigger than zero, but smaller than any positive real number. The philosopher Berkeley called such infinitesimals "ghosts of recently departed quantities."

The textbook is of the opinion that the first, rigorous, approach is the only way to go. Personally, I find the second approach extremely useful, even if it is just "walking on clouds." I also think you need to see it, since if you should need calculus in applied science (physics, chemistry, atmospheric chemistry,...) this is most likely the language you will see. But I'm honestly not sure which approach is less confusing to see first, so I'm going to accept the following wisdom: When in doubt, follow the textbook. More or less.

2 Defining instantaneous velocity

As said above, the essence of motion is changing from one position to another. So, let's suppose that an object changes its position over time. As Zeno pointed out, at any given time, it has only one position. Thus, if we denote the object's position by x, then x is a function of the time t: there exists a function f such that

$$x = f(t).$$

Consider what happens near a fixed time t_0. As a small amount of time elapses, the object's position changes by a small amount; the velocity is the change in position divided by the change in time. For some reason, it is customary to use the Greek letter Δ (capital delta) to represent "change in." Thus, with this notation, the above sentence states that

$$\text{velocity} = \frac{\Delta x}{\Delta t}.$$

There's a bit of a problem here, though. If we specifically want the velocity *at the instant* t_0, then we don't have any change in time to work with: $\Delta t = 0$. Likewise, within the single instant, there is no change in position: $\Delta x = 0$. So, the expression above would tell us that velocity $= 0/0$. Since $0/0$ is undefined, this is not terribly helpful.

However, we have been studying a way to "fill in" such undefined values: use limits. Thus, we define the *instantaneous velocity at* t_0, denoted $dx/dt|_{t=t_0}$, to be

$$\left.\frac{dx}{dt}\right|_{t=t_0} = \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta t},$$

provided that this limit exists. The notation follows the convention that "when you take a limit, you should replace Greek letters by Roman letters." In this case, we replace the Greek letter Δ by the Roman letter d.

Recall that the object starts at time t_0. If the time changes by $\Delta t = h$, then the corresponding change in position is $\Delta x = f(t_0+h) - f(t_0)$. Thus, the above equation can also be written

$$\left.\frac{dx}{dt}\right|_{t=t_0} = \lim_{h \to 0} \frac{f(t_0+h) - f(t_0)}{h}.$$

Finally, if we bring to bear all of the different notations we're likely to use for this, we'll get

$$f'(t_0) = \frac{df}{dt}(t_0) = \left.\frac{dx}{dt}\right|_{t=t_0} = \lim_{h \to 0} \frac{f(t_0+h) - f(t_0)}{h}.$$

This quantity (when it exists) is called the *derivative of f at t_0*.

3 A few more bits on continuity

People generally learn stuff better if they see it over a period of time, rather than all at once. Thus, I've decided to distribute the subject of continuity over multiple lectures.[1]

Theorem. Every polynomial or rational function is continuous on its natural domain. The same holds if you throw in n^{th} roots.

The proof of this theorem is by repeated applications of the Main Limit Theorem. I won't try to give the complete proof, but I will give you an example.

Example. Show, using the Main Limit Theorem, that the function f defined by

$$f(x) = \frac{1+\sqrt{2x}}{x^3-13}$$

is continuous.

Solution. For every real number c such that $f(c)$ is defined, we have

$$\begin{aligned}
\lim_{x\to c} f(x) &= \lim_{x\to c} \frac{1+\sqrt{2x}}{x^3-13} \\
&= \frac{\lim_{x\to c} 1+\sqrt{2x}}{\lim_{x\to c} x^3-13} \\
&= \frac{1+\sqrt{\lim_{x\to c} 2x}}{c^3-13} \\
&= \frac{1+\sqrt{2c}}{c^3-13} = f(c).
\end{aligned}$$

By hypothesis, $f(c)$ is defined, i.e., the last line makes sense. By the Main Limit Theorem, the previous line makes sense and is equal to it, and so on all the way up. Thus, for every c in the domain of f,

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

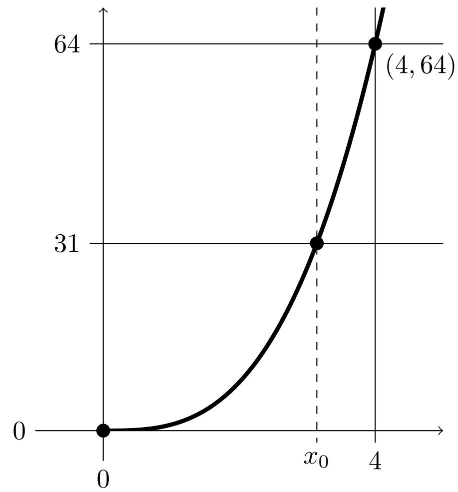
In other words, f is continuous. $\square$

Finally, an example of using the Intermediate Value Theorem that hearkens back to the bonus problem from the first lecture:

Example. Show that $\sqrt[3]{31}$ exists. In other words, show that there is a positive real number x_0 such that $x_0^3 = 31$.

[1] Translation: I did not really get to say all I wanted on continuity last lecture, so I'm trying to cram in a few extra notes at the end of this one.

Solution.



Let f be the function defined by $f(x) = x^3$. Since f is a polynomial function, f is continuous on its domain $(-\infty, \infty)$, and in particular on the interval $[0, 4]$. Observe that

$$f(0) = 0 < 31 < 64 = f(4).$$

Hence, there exists some x_0 such that $0 < x_0 < 4$ and $f(x_0) = 31$. $\square$

Assignment 12 (due Monday, October 31, a.k.a. Halloween)

Section 1.3, Problems 3, 4, 7, and 8. Remember, these problems are about showing you understand how to use the Main Limit Theorem, not about finding the limits. Problems 4 and 8 will be graded carefully.

Section 1.6, Problems 2, 4, 6-8, 32, and 33. Problems 6-8 and 32 will be graded carefully. You do not need to give ε-δ proofs for these problems.

Bonus problem: Let f be a function (which you do not get to choose). Consider the statement

> For all x_0 in the domain of f, for all $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$.

Explain why this is equivalent to the statement that "f is continuous." (One dilemma you may need to address: Why does it make no difference if we write $|x - x_0| < \delta$ rather than $0 < |x - x_0| - \delta$?)

Assignment 13 (due Wednesday, 2 November)

NOTE: From this assignment on, you no longer need to write anything about "By the Main Limit Theorem,..." when showing your work to take a limit. (You should, however, continue to show your work.)

Use the Intermediate Value Theorem to prove that, no matter what Diophantus[2] of Alexandria might have thought, $\sqrt{2}$ does, in fact, exist. (In other words, there exists a positive real number x_0 such that $x_0^2 = 2$.) This problem will be graded carefully.

Section 2.2, Problems 45-48 and 51, 52. Be sure to follow the instructions carefully on 51 and 52; these problems are as much about *how* you find the derivative, as what answer you get. Problems 46, 48, and 52 will be graded carefully.

Let $a \neq 0$ be a real number (which you don't get to choose). Let f be the function defined by

$$f(x) = ax.$$

Show that f is continuous using an ε-δ proof. This problem will be graded carefully.

[2] See Lecture 3, page 2.

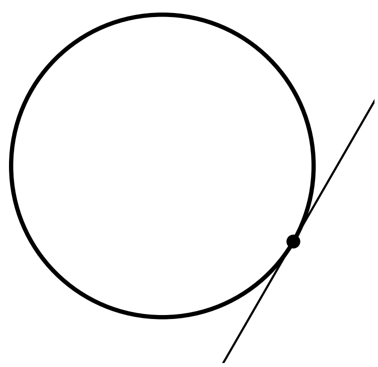
Math 131, Lecture 15

Charles Staats

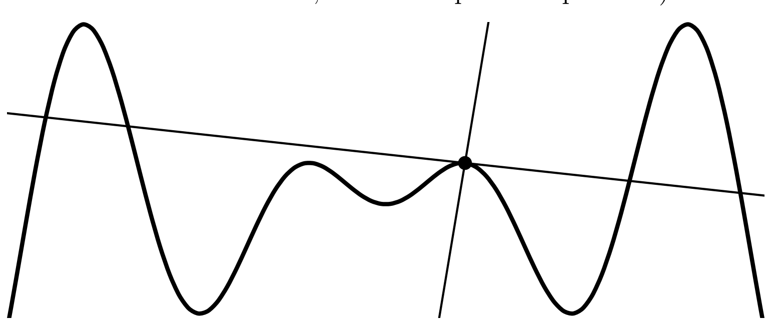
Monday, 31 October 2011

1 The derivative as the slope of the tangent line

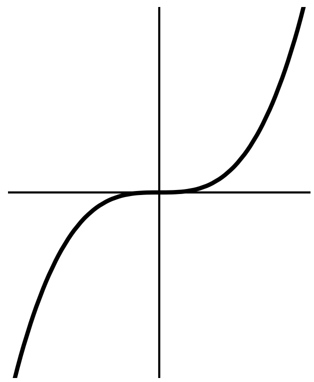
In classical geometry, the tangent to a curve was the line that somehow "touched the curve without crossing it." Euclid attempted to make this precise by describing the tangent as the line that intersected the curve in only one point. His definition works quite well for circles (and also ellipses, parabolas, and hyperbolas):



However, it can fail rather drastically for more complicated curves. In the curve below, the almost-vertical line is the one that intersects the curve in only one point, while the almost-horizontal line clearly "ought" to be the tangent line. (Intuitively, the almost-vertical line crosses the curve, while the almost-horizontal line does not—at least, not at the point in question.)



For another example, in the following picture, neither the vertical nor the horizontal line really "touches the curve without crossing it." Each of them intersects the curve exactly once. But if one of them is the tangent line, it is the horizontal line rather than the vertical line.



Thus, we take another approach to defining what exactly the tangent line should be. An easier definition is to define a secant line—that is, a line that passes through two specified points on a curve. This is easy to specify, since two points determine a line. We want to think of a tangent line as a "secant line that passes through the same point twice." Unfortunately, this does not actually make any sense.

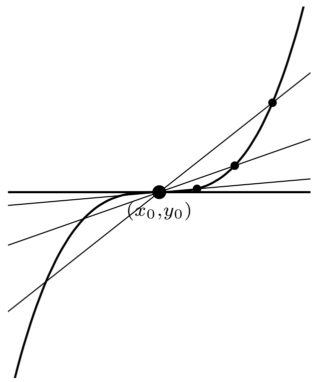
To remedy the situation, we consider another way of specifying a line: a point (x_0, y_0) together with a slope $\Delta y/\Delta x$. Thus, the secant line through (x_0, y_0) and (x, y) is the line passing through (x_0, y_0) with slope equal to

$$\frac{\Delta y}{\Delta x} = \frac{y - y_0}{x - x_0}.$$

If we want to take the tangent line at (x_0, y_0), we already have a point through which the line should pass. We just need to know what its slope ought to be. This is essentially the same problem we were faced with last lecture—we need a definition for "slope *at a point*," in spite of the fact that slope is, inherently, a property relating two different points. And we solve it the same way: we take a limit. We say that the slope of the tangent line is

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

The picture below shows the tangent line as a limit of secant lines:



Now, if you recall the previous definition of the derivative, you will see that, if y is given as $y = f(x)$ for some function f, then in fact, we will have the slope of the tangent line equal to the derivative:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \left.\frac{dy}{dx}\right|_{x=x_0} = f'(x_0).$$

This gives us the following

Definition. Let f be a function defined at x_0. The *tangent line* to f at x_0 is the line passing through the point $(x_0, f(x_0))$ and having slope equal to $f'(x_0)$, provided that this derivative exists.

Let's do an example.

Example. Let the function f be defined by

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

Compute the derivative of f at $x_0 = 1$. Plot the function and the line tangent to f at x_0.

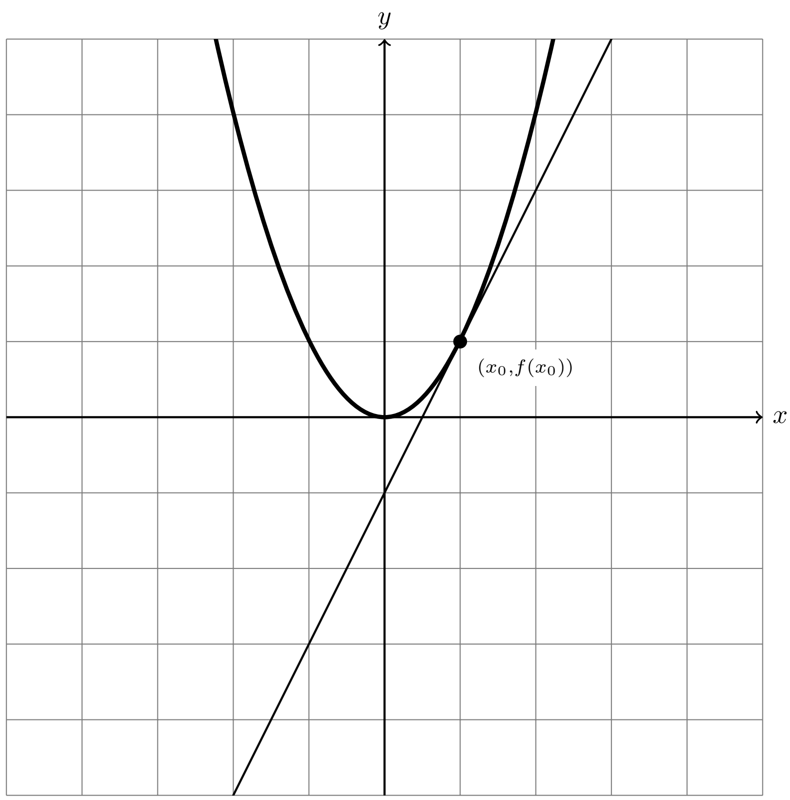
Solution. First, let's solve for Δy in terms of Δx:

$$\begin{aligned}\Delta y &= f(x_0 + \Delta x) - f(x_0)\\ &= (1 + \Delta x)^2 - 1^2\\ &= 1 + 2\Delta x + (\Delta x)^2 - 1\\ &= 2\Delta x + (\Delta x)^2.\end{aligned}$$

Thus, we have

$$\begin{aligned}
f'(x_0) &= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{2\Delta x + (\Delta x)^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} 2 + \Delta x \\
&= 2.
\end{aligned}$$

Now, we plot the function $y = f(x)$, together with line passing through $(x_0, f(x_0)) = (1, 1)$ and having slope $f'(x_0) = 2$:



$\square$

2 Infinitesimals

The idea of infinitesimals, as it relates to slopes of tangent lines, is to define the tangent line to f at x_0 as the line through (x_0, y_0) and another point $(x_0 + dx, y_0 + dy)$ that is "infinitely close" to the first point. What this means, in this example, is that dx is "so small" that $dx^2 = 0$, even though dx is not zero.

This sort of makes sense, in that the square of a small number is a much smaller number; for instance,

$$0.001^2 = 0.000001$$

It does not really make sense—no nonzero number can square to zero—but that's why I called this "walking on clouds."

To start with, we treat the "infinitesimal changes" dx and dy exactly as though they were more conventional changes Δx and Δy. Our earlier computation of Δy in terms of Δx still holds:

$$\begin{aligned}
dy &= 2dx + dx^2 \\
dy &= 2dx && \text{since } dx^2 = 0 \\
\frac{dy}{dx} &= 2
\end{aligned}$$

when evaluated at the point $x_0 = 1$.

Assignment 13 (due Wednesday, 2 November)

NOTE: From this assignment on, you no longer need to write anything about "By the Main Limit Theorem,..." when showing your work to take a limit. (You should, however, continue to show your work.)

Use the Intermediate Value Theorem to prove that, no matter what Diophantus[1] of Alexandria might have thought, $\sqrt{2}$ does, in fact, exist. (In other words, there exists a positive real number x_0 such that $x_0^2 = 2$.) This problem will be graded carefully.

Section 2.2, Problems 45-48 and 51, 52. Be sure to follow the instructions carefully on 51 and 52; these problems are as much about *how* you find the derivative, as what answer you get. Problems 46, 48, and 52 will be graded carefully.

Let $a \neq 0$ be a real number (which you don't get to choose). Let f be the function defined by

$$f(x) = ax.$$

Show that f is continuous using an ε-δ proof. This problem will be graded carefully.

Assignment 14 (due Friday, 4 November)

Section 2.2, Problems 37–44. The even-numbered problems will be graded carefully.

For each of Problems 1–4 in Section 2.2, do the following steps:

(a) Find the indicated derivative using infinitesimals.

(b) Find the indicated derivative using the limit definition.

(c) Graph the function together with the tangent line at the indicated point.

Problems 2 and 4 will be graded carefully.

[1] See Lecture 3, page 2.

Math 131, Lecture 16

Charles Staats

Wednesday, 1 November 2011

1 Some alternate ways to state the limit definition of the derivative

For this section, we're going to use the notation $f'(x_0)$ rather than $dy/dx|_{x=x_0}$. Our basic definition of the derivative has been

$$\boxed{f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.} \tag{1}$$

One key to mastering mathematics is being able to move facilely among different ways of saying the same thing; which way you want to say it may depend on what you want to use it for. We're going to review some other ways to write the definition of the derivative, using the various relations among $x, x_0, \Delta x, y, \Delta y, f(x_0), \ldots$.

First, observe that

$$\begin{aligned} \Delta y &= y - y_0 = f(x) - f(x_0) \qquad \text{and} \\ \Delta x &= x - x_0. \end{aligned}$$

Thus,

$$\frac{\Delta y}{\Delta x} = \frac{f(x) - f(x_0)}{x - x_0},$$

and

$$\begin{aligned} &\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \ell \\ &\iff \forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.t. if } 0 < |\Delta x - 0| < \delta, \text{ then } \left| \frac{\Delta y}{\Delta x} - \ell \right| < \varepsilon \\ &\iff \forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.t. if } 0 < |x - x_0| < \delta, \text{ then } \left| \frac{f(x) - f(x_0)}{x - x_0} - \ell \right| < \varepsilon \\ &\iff \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \ell. \end{aligned}$$

In other words, an alternate definition for the derivative is given by

$$\boxed{f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.} \tag{2}$$

This definition highlights the feature that the derivative only depends on what is happening to f near x_0. If we look at a different function g that cannot be distinguished from f near x_0, then f and g will have the same derivative at x_0; i.e., $f'(x_0) = g'(x_0)$.

Another way to state the definition of the derivative is to express Δy in terms of x_0 and Δx, rather than x_0 and x.

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

since $x = x_0 + \Delta x$. Thus, we have

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Making the traditional change of notation $\Delta x = h$, we find that

$$\boxed{f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.} \tag{3}$$

The expression inside the limit is the infamous “difference quotient.”

2 Derivative as a function

In the definition of (3), one feature is that there are no appearances of the letter x except in the variable x_0. Thus, we can rename x_0 as x, obtaining

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

The interesting feature here is that when we rewrite the definition this way, it becomes obvious that we have defined more than a number $f'(x_0)$; we have defined a *function* f'.

There’s a subtlety here that confused me when I first saw this sort of thing. It involves the interplay of intuition and rigorous mathematics. Intuitively, when we write x, we think of it as a variable—something that is allowed to range over many different numbers. On the other hand, when we write x_0, we think of this as a particular value of x, a particular number; we just don’t happen to know what number it is. These intuitions are valuable. However, it is equally valuable to realize that these intuitions have absolutely no reflection in the rigorous mathematics. As far as the pure logic is concerned, x and x_0 are both variables, and that’s all there is to it. So whenever we have a statement that involves only one, we can substitute the other, and get an equally true expression that feels very different, intuitively.

This is typical of a certain kind of reasoning that appears sometimes in mathematics. First, you let your intuition guide you, as we did (more or less) in defining the derivative. Then you do something with rigorous mathematics to change the statement into something equivalent, but that feels intuitively very different. At this point, you may feel like your head wants to explode: your intuition is screaming that what you've done can't possibly be right, but you can't see any flaws in your logic. It may be tempting to give up and think about something else. But instead, you may force yourself to stay on task, to turn the thing over and over in your head until you either find a flaw in the logic, or find a way of thinking about it that your intuition will accept. Depending on the difficulty of the thing in question, resolving the conflict may take moments, hours, days, weeks, months, or years. But the longer you spend puzzling over it, the greater will be your feeling of enlightenment when it finally "clicks."

On the other hand, some of you may be thinking that it was obvious that the derivative is a function. You may even feel a bit smug about the fact that this "revelation" was clear to you from the beginning. Perhaps you should. But I think it is more likely that you were not following my lectures closely, but were instead thinking about the derivative in terms you have learned in the past. Or perhaps you never really understood the intuition of x_0 as a "fixed value we don't know," versus x as a "variable." Either way, I suggest you review the previous buildup to the definition of the derivative. Try to understand with your whole mind—both logic and intuition. If you succeed, you may get a part of the revelatory moment that you will otherwise have been cheated of.

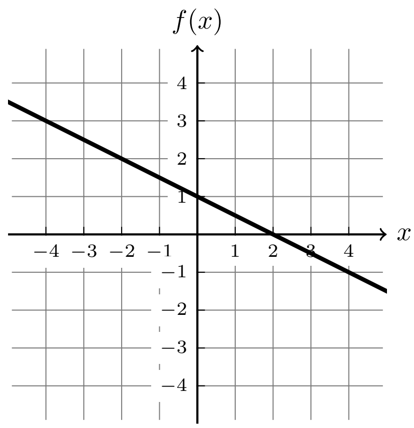
Now, enough philosophizing. Since we've established that the derivative f' is a function, there are two obvious sorts of questions:

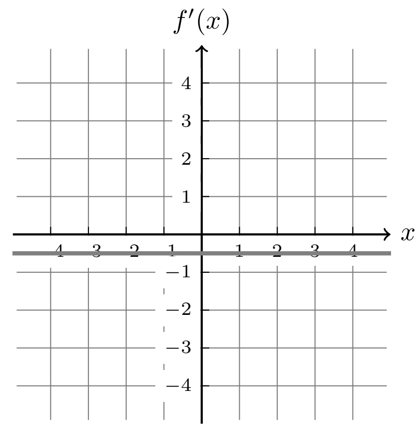
1. How do we find a formula for the function, if one exists?

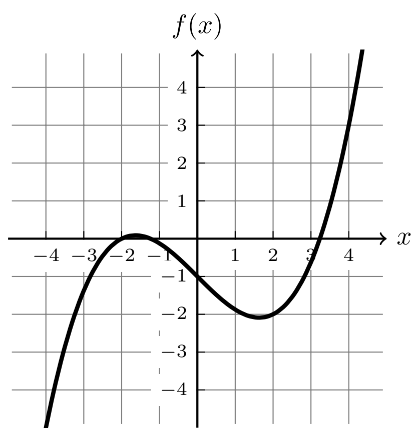
2. How do we characterize the function, even if it does not have a formula we can write down?

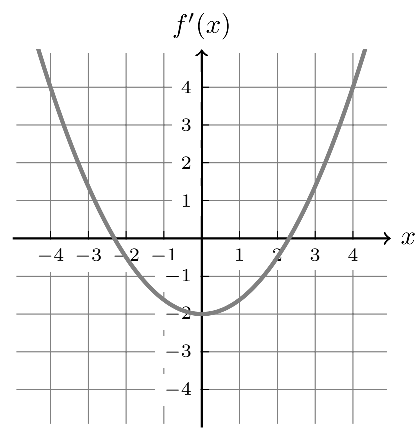
We'll spend a lot of time on both of these, but in light of the homework I've assigned you for Friday, I'm going to spend the rest of this lecture on a version of the second problem. Specifically: If someone gives you a graph of the function, how do you graph its derivative? We'll approach this mainly through examples. My plan (which I may or may not have time for) is to give you a few minutes to try the following examples on your own, and then we will go over them together.

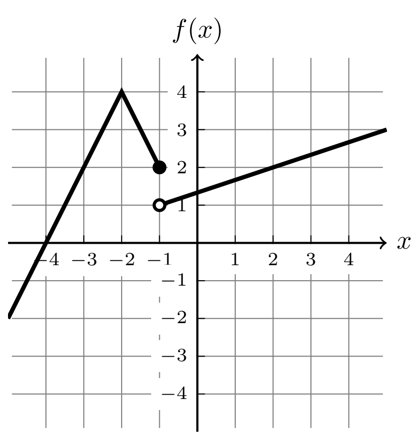
Example. The graph of a function f is given on the left. On the right, sketch the graph of the function f'. Remember: above each point x on the x-axis, the value of f' should be the slope of the tangent line to f at x. If f does not have a unique tangent line at x, then $f'(x)$ will not exist.

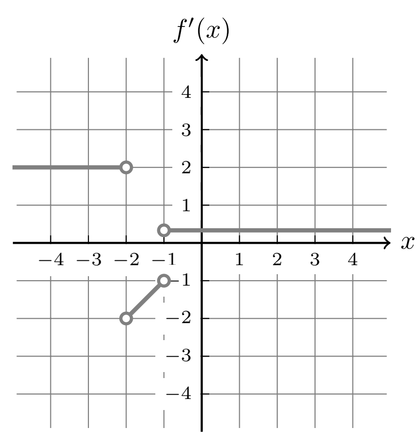












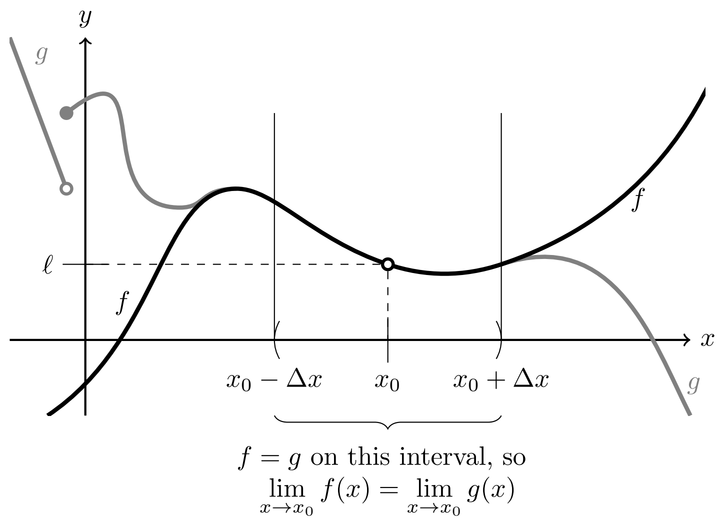
3 Local nature of the limit (and derivative)

I'm probably not going to have time to really go over this section in the lecture, but I would feel like I would not be fulfilling my responsibilities as a Math 131 teacher if I did not at least mention it in the lecture notes.

Recall that, in the most vague terms, the statement

$$\lim_{x \to x_0} f(x) = \ell$$

means something like "when x is near x_0, then $f(x)$ is near ℓ." Thus, it seems like this limit should only depend on "what f is doing near x_0." In particular, it should only depend on how f behaves on an interval $(x_0 - \Delta x, x_0 + \Delta x)$.



The way we say that the limit "only depends on what f is doing near x_0" is that if we replace f by a different function g that "looks the same near x_0," then we are guaranteed to get the same answer. More precisely, we have the following theorem:

Theorem. Suppose that f and g are two functions. Let Δx be positive. If f and g are defined and agree on the interval $(x_0 - \Delta x, x_0 + \Delta x)$, then

$$\lim_{x \to x_0} f(x) \text{ exists if and only if } \lim_{x \to x_0} g(x) \text{ exists.}$$

Moreover, if the two limits exist, then they are equal.

Proof. Assume that

$$\lim_{x \to x_0} f(x) = \ell.$$

We will then show that $\lim_{x \to x_0} g(x) = \ell$.

Let $\varepsilon > 0$ be given.

Since $\lim_{x\to x_0} f(x) = \ell$, there exists $\delta_1 > 0$ such that if $0 < |x - x_0| < \delta_1$, then $|f(x) - \ell| < \varepsilon$. Set $\delta = \min\{\delta_1, \Delta x\}$.

Assume $0 < |x - x_0| < \delta$. Since $|x - x_0| < \delta \leq \Delta x$, we know $f(x) = g(x)$. Consequently,

$$\begin{aligned} |g(x) - \ell| &= |f(x) - \ell| \\ &< \varepsilon, \end{aligned}$$

since $0 < |x - x_0| < \delta \leq \delta_1$.

Therefore,

$$\lim_{x\to x_0} g(x) = \ell,$$

as claimed.

Similar reasoning shows that, if

$$\lim_{x\to x_0} g(x) = \ell,$$

then $\lim_{x\to x_0} f(x) = \ell$. □

Since one definition for the derivative is

$$\left.\frac{dy}{dx}\right|_{x=x_0} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

the theorem tells us that the derivative of f at x_0 depends only on how f behaves near x_0.

Assignment 14 (due Friday, 4 November)

Section 2.2, Problems 37–44. The even-numbered problems will be graded carefully.

For each of Problems 1–4 in Section 2.2, do the following steps:

(a) Find the indicated derivative using infinitesimals.

(b) Find the indicated derivative using the limit definition.

(c) Graph the function together with the tangent line at the indicated point.

Problems 2 and 4 will be graded carefully.

Assignment 15 (due Monday, 7 November)

Section 2.2, Problems 5–8 and 51–54. Follow the instructions. Remember, these problems are more about *how* you find the derivative than what derivative you find. The even-numbered problems will be graded carefully.

Section 2.3, Problems 1-4. These will not be graded carefully.

Draw a picture that explains why the difference quotient

$$\frac{f(x+h)-f(x)}{h}$$

gives the slope of a secant line to the curve $y = f(x)$. Hint: your intuition may like this problem better if you think in terms of x_0 rather than x. This problem will be graded carefully.

Math 131, Lecture 17

Charles Staats

Friday, 4 November 2011

1 Some example computations

We're going to do compute some derivatives as functions using the definition

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

Example. Suppose that $f(x) = x$. Compute a formula for the function f'.

Solution.

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h\to 0} \frac{(x+h) - x}{h} \\ &= \lim_{h\to 0} \frac{h}{h} \\ &= \lim_{h\to 0} 1 \\ &= 1. \end{aligned}$$

□

Example. Suppose that $f(x) = mx + b$. Since $y = f(x)$ is a line, the tangent line will be the line itself; its slope, of course, is m. Thus, we may suppose that $f'(x) = m$ for all x. Prove this using the limit definition.

Solution.

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \frac{m(x+h)+b-(mx+b)}{h} \\
&= \lim_{h\to 0} \frac{mx+mh+b-mx-b}{h} \\
&= \lim_{h\to 0} \frac{mh}{h} \\
&= \lim_{h\to 0} m \\
&= m.
\end{aligned}$$

□

Example. Let f be the function defined by $f(x) = x^2 + x - 3$. Compute a formula for the function f'.

Solution.

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \frac{(x+h)^2+(x+h)-3-x^2-x+3}{h} \\
&= \lim_{h\to 0} \frac{x^2+2xh+h^2+x+h-3-x^2-x+3}{h} \\
&= \lim_{h\to 0} \frac{2xh+h^2+h}{h} \\
&= \lim_{h\to 0} \frac{(2x+h+1)h}{h} \\
&= \lim_{h\to 0} 2x+h+1 \\
&= 2x+1.
\end{aligned}$$

□

Example. Let f be the function defined by $f(x) = 1/x$. Compute a formula for the derivative of f (except at $x = 0$, of course).

Solution.

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \frac{\frac{1}{x+h}-\frac{1}{x}}{h} \cdot \frac{x(x+h)}{x(x+h)} \\
&= \lim_{h\to 0} \frac{x-(x+h)}{hx(x+h)} \\
&= \lim_{h\to 0} \frac{-h}{hx(x+h)} \\
&= \lim_{h\to 0} \frac{-1}{x(x+h)} \\
&= \frac{-1}{x^2}.
\end{aligned}$$

□

Notation. It can be rather tiresome to write, for instance, "the derivative of the function f defined by $f(x) = x^2 + x - 3$." In the future, we will sometimes abbreviate this by

$$\frac{d}{dx}(x^2 + x - 3).$$

2 Product rule

Suppose that we have u and v, two functions of x. Suppose we know how to calculate the derivatives du/dx and dv/dx. We can use this to calculate the derivative of the product $u \cdot v$, by means of the product rule.

Warning. *It may be tempting to write that*

$$\frac{d}{dx}(u \cdot v) = \frac{du}{dx} \cdot \frac{dv}{dx}.$$

THIS IS NOT TRUE. *For instance, suppose $u(x) = 2$ and $v(x) = x$. Then*

$$\frac{d}{dx}(u \cdot v) = \frac{d}{dx}(2 \cdot x) = 2,$$

since $y = 2x$ is a line of slope 2. However, the "naive product rule" would give us

$$\frac{d}{dx}(2 \cdot x) = \frac{d}{dx}(2) \cdot \frac{d}{dx}(x) = 0 \cdot 1 = 0.$$

The naive product rule gives the wrong answer.

Leibniz gave a cute derivation of the product rule using infinitesimals. The first equation in this proof may seem a bit confusing at first; I'll explain it

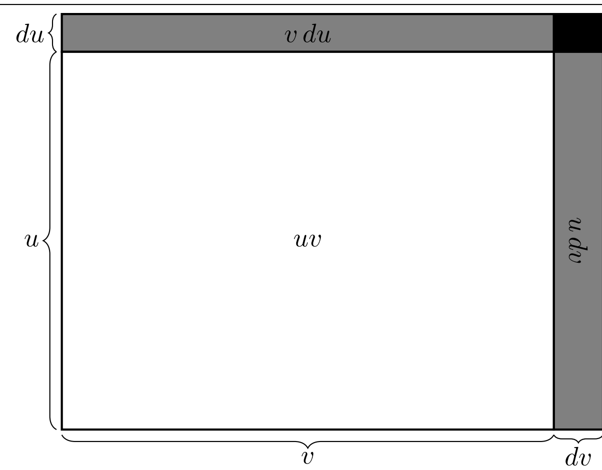


Figure 1: A visual illustration of the product rule. The area of the white rectangle is uv; the area of the total rectangle is $(u+du)(v+dv)$; and the change in area, $d(uv)$, is their difference. The black rectangle, with area $du\,dv$, is so small that its contribution is "can be neglected."

afterwards, but if I give it now, the proof will not seem so "cute." Remember, the key "fact" about infinitesimals is that if you multiply two of them together, you get something "doubly infinitesimal," which we typically consider equal to zero. In particular, $du\,dv = 0$.

$$\begin{aligned} d(uv) &= (u+du)(v+dv) - uv \\ &= uv + u\,dv + v\,du + du\,dv - uv \\ &= u\,dv + v\,du. \end{aligned}$$

Dividing through by dx, we see that

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}.$$

Now, the promised explanation of the first line: we have two functions u and v of x. But we really have three functions: the one we care about is the function f defined by $f = uv$, i.e.,

$$f(x) = u(x) \cdot v(x).$$

Thus,

$$\begin{aligned} df &= f(x+dx) - f(x) \\ &= u(x+dx)v(x+dx) - u(x)v(x). \end{aligned}$$

Recall that

$$\begin{aligned} du &= u(x+dx) - u(x), \qquad \text{hence} \\ u + du &= u(x+dx). \end{aligned}$$

Similarly, $v + dv = v(x + dx)$, and so we have

$$\begin{aligned} df &= u(x+dx)v(x+dx) - u(x)v(x) \\ &= (u+du)(v+dv) - uv. \end{aligned}$$

(By an abuse of notation, we're writing things like u for $u(x)$ when it suits us to do so.)

Example. Use the product rule to find (in this order) the derivatives of x^2, x^3, and x^4 with respect to x.

Solution.

$$\begin{aligned} \frac{d}{dx}x^2 &= \frac{d}{dx}(x \cdot x) \\ &= x\frac{dx}{dx} + x\frac{dx}{dx} \\ &= x + x \\ &= 2x. \end{aligned}$$

$$\begin{aligned} \frac{d}{dx}x^3 &= \frac{d}{dx}(x \cdot x^2) \\ &= x\frac{d}{dx}(x^2) + x^2\frac{d}{dx}(x) \end{aligned}$$

We just calculated $\frac{d}{dx}(x^2) = 2x$, so this is equal to

$$\begin{aligned} &= x \cdot 2x + x^2 \cdot 1 \\ &= 2x^2 + x^2 \\ &= 3x^2. \end{aligned}$$

$$\begin{aligned} \frac{d}{dx}x^4 &= \frac{d}{dx}(x \cdot x^3) \\ &= x \cdot \frac{d}{dx}(x^3) + x^3\frac{d}{dx}(x) \\ &= x \cdot 3x^2 + x^3 \cdot 1 \\ &= 3x^3 + x^3 \\ &= 4x^3. \end{aligned}$$

□

You may start to notice a pattern here. This pattern will continue: if we calculate on out to $\frac{d}{dx}x^{n-1}$, we'll find that it is equal to $(n-1)x^{n-2}$. Using this fact, we find that

$$\begin{aligned}\frac{d}{dx}x^n &= \frac{d}{dx}(x \cdot x^{n-1}) \\ &= x \cdot \frac{d}{dx}(x^{n-1}) + x^{n-1}\frac{d}{dx}(x) \\ &= x \cdot (n-1)x^{n-2} + x^{n-1} \cdot 1 \\ &= (n-1)x^{n-1} + x^{n-1} \\ &= nx^{n-1},\end{aligned}$$

so the pattern always keeps going. (This is a version of "proof by induction.")

Assignment 15 (due Monday, 7 November)

Section 2.2, Problems 5–8 and 51–54. Follow the instructions. Remember, these problems are more about *how* you find the derivative than what derivative you find. The even-numbered problems will be graded carefully.

Section 2.3, Problems 1-4. These will not be graded carefully.

Draw a picture that explains why the difference quotient

$$\frac{f(x+h)-f(x)}{h}$$

gives the slope of a secant line to the curve $y = f(x)$. Hint: your intuition may like this problem better if you think in terms of x_0 rather than x. This problem will be graded carefully.

Assignment 16 (due Wednesday, 9 November)

Section 2.2, Problems 11–14 and 55–58. Follow the instructions. Remember, these problems (except for 57 and 58) are more about *how* you find the derivative than what derivative you find. The even-numbered problems will be graded carefully (although a certain amount of leeway will be provided on problem 58).

Section 2.3, Problems 11–14 and 23–26. (Hint: 23–26 are easier if you use the product rule.) The even-numbered problems will be graded carefully.

Translate the statement

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0)$$

into ε-δ language. (Hint: when you see $f'(x_0)$, treat it like ℓ. Also, treat $\Delta y/\Delta x$ as a function of Δx.) Then, use the resulting statement to prove the following:

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. if } |\Delta x| < \delta, \text{ then } \Delta y \text{ is within } \varepsilon \text{ of } f'(x_0)\Delta x.$$

You will need to handle $\Delta x = 0$ as a separate case. This statement is a rigorous version of the statement that "When Δx is small, then Δy is approximately $\frac{dy}{dx}\Delta x$."

Test II around Wednesday, 16 November

Experienced teachers of Math 131 tell me that you will probably have a lot of papers and the like due around the time of the test, and consequently will not have a lot of time to study for it. Thus, I suggest you start studying now. You may also want to think in terms of "practicing" rather than "studying": redoing old quiz and homework problems (without looking at the solutions, if you have them, until afterwards) may be more helpful than simply reading over them.

Math 131, Lecture 18: Rules for differentiation

Charles Staats

Monday, 7 November 2011

The process of taking a derivative, often called "differentiation," is extremely important. Moreover, unlike many important things in mathematics, differentiation is actually possible to do. Any time you have a function given by a formula, the rules in this lecture will allow you to find its derivative.

These rules need to be memorized. Ideally, they should become so ingrained that you can use them without having to think about them.

1 The "easy rules"

There are a few "easy" rules for differentiation.

Theorem. (Constant rule) If f is the function defined by $f(x) = c$, where c is a (constant) real number, then $f'(x) = 0$ for all x.

Proof. This is a special case of the $mx + b$ rule we proved last time, but let's do it again anyway.

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h\to 0} \frac{c - c}{h} \\ &= \lim_{h\to 0} \frac{0}{h} \\ &= 0. \end{aligned}$$

□

Theorem. (Sum rule) If f and g are differentiable functions, then

$$(f + g)' = f' + g'.$$

In words, "the derivative of a sum is the sum of the derivatives."

Proof. We apply one of the limit definitions of the derivative:

$$\begin{aligned}
(f+g)'(x) &= \lim_{h\to 0} \frac{(f+g)(x+h)-(f+g)(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)+g(x+h)-f(x)-g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} + \frac{g(x+h)-g(x)}{h} \\
&= f'(x)+g'(x) \\
&= (f'+g')(x).
\end{aligned}$$

Since this holds for all x at which f and g are defined, we have the equality of functions

$$(f+g)' = f' + g'. \qquad \square$$

Theorem. (Multiplication by a constant) If f is a differentiable function of x and c is a (constant) real number, then

$$\frac{d}{dx}(cf(x)) = c\frac{df}{dx}.$$

Proof. This is, again, a special case of the $mx+b$ thing. This time, we're going to derive it from the product rule.

$$\begin{aligned}
\frac{d}{dx}(cf(x)) &= c\frac{df}{dx} + f(x)\frac{d}{dx}(c) \\
&= c\frac{df}{dx} + f(x)\cdot 0 && \text{(constant rule)} \\
&= c\frac{df}{dx}.
\end{aligned} \qquad \square$$

Theorem. (Difference rule) If f and g are differentiable functions, then $(f-g)' = f' - g'$.

Proof.

$$\begin{aligned}
(f-g)' &= \big(f + (-1)\cdot g\big)' \\
&= f' + \big((-1)g\big)' && \text{(sum rule)} \\
&= f' + (-1)g' && \text{(multiplication by a constant)} \\
&= f' - g'.
\end{aligned} \qquad \square$$

2 The power rule; polynomials

The power rule is fairly easy, but a bit less intuitive than the "easy rules." It was mentioned briefly at the end of the last lecture.

Theorem. When n is a positive integer,

$$\frac{d}{dx}x^n = nx^{n-1}.$$

(Actually, this theorem applies whenever n is a real number, but we won't be able to prove that for some time.)

To understand the proof of the Power Rule, we need a technique called *mathematical induction*. Suppose we have a condition $P(n)$ on n. The "induction principle" says that to show $P(n)$ is true whenever n is a positive integer, we can do show the following:

1. $P(1)$ is true.
2. Whenever $P(n)$ is true, then $P(n+1)$ is also true.

Thus, $P(1)$ is true; since $P(1)$ is true, $P(2)$ is also true; since $P(2)$ is true, $P(3)$ is also true; and so on.

One standard metaphor here is that in step 2, we set up a chain of dominoes; in step 1, we knock over the first one, which then knocks over the second one, which then knocks over the third one, etc.

Proof. Let $P(n)$ be the statement that $D_x(x^n) = nx^{n-1}$; this is a condition on n.

1. We first show that $P(1)$ is true, i.e., that $D_x(x) = 1$:

$$\begin{aligned}\frac{d}{dx}(x) &= \lim_{h\to 0}\frac{(x+h)-h}{h}\\ &= \lim_{h\to 0}\frac{h}{h}\\ &= 1,\end{aligned}$$

as desired.

2. We now show, using the product rule, that whenever $P(n)$ is true, then $P(n+1)$ is also true.

$$\begin{aligned}\frac{d}{dx}x^{n+1} &= \frac{d}{dx}(x\cdot x^n)\\ &= x\frac{d}{dx}(x^n) + x^n\frac{d}{dx}(x) && \text{(product rule)}\\ &= x\cdot nx^{n-1} + x^n\cdot 1 && \text{(since } P(n) \text{ is true)}\\ &= nx^n + x^n\\ &= (n+1)x^n.\end{aligned}$$

Thus, by induction, $P(n)$ is true for every positive integer n. In other words, for every positive integer n,

$$\frac{d}{dx}x^n = nx^{n-1}. \qquad \square$$

Using the power rule, together with the "easy rules," we can, in principle, compute the derivative of any polynomial.

Example 1. Differentiate $x^2 - 4x + 1$.

Solution.

$$\begin{aligned}
\frac{d}{dx}(x^2 - 4x + 1) &= \frac{d}{dx}(x^2) - \frac{d}{dx}(4x) + \frac{d}{dx}(1) && \text{(sum rule)} \\
&= \frac{d}{dx}(x^2) - 4\frac{d}{dx}(x) + 0 && \text{(constant multiple; constant)} \\
&= 2x - 4 \cdot 1 + 0 && \text{(power rule)} \\
&= 2x - 4. && \square
\end{aligned}$$

Example 2. Differentiate $2x^3 - \frac{1}{2}x^2 - x + \frac{17246}{937}$.

Solution.

$$\begin{aligned}
\frac{d}{dx}\left(2x^3 - \tfrac{1}{2}x^2 - x + \tfrac{17246}{937}\right) &= 2 \cdot 3x^2 - \tfrac{1}{2} \cdot 2x - 1 + 0 \\
&= 6x^2 - x - 1. \qquad \square
\end{aligned}$$

3 Proof of the product rule

Recall the product rule,

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx},$$

and the (non-rigorous) infinitesimal derivation:

$$\begin{aligned}
d(uv) &= (u + du)(v + dv) - uv \\
&= uv + u\,dv + v\,du + du\,dv - uv \\
&= u\,dv + v\,du \\
\frac{d}{dx}(uv) &= u\frac{dv}{dx} + v\frac{du}{dx}.
\end{aligned}$$

We are now going to show how to prove the product rule rigorously. Pay attention to how what we are doing rigorously corresponds to the non-rigorous infinitesimal method.

Theorem. Suppose that u is a function of x such that $du/dx|_{x=x_0}$ exists. Likewise, suppose that v is a function of x such that $dv/dx|_{x=x_0}$ exists. Then the derivative of the product uv at x_0 exists, and

$$\frac{d}{dx}(uv)\bigg|_{x=x_0} = u_0\,\frac{dv}{dx}\bigg|_{x=x_0} + v_0\,\frac{du}{dx}\bigg|_{x=x_0}.$$

Proof. First, we write $\Delta(uv)$ in terms of Δu and Δv:

$$\begin{aligned}\Delta(uv) &= uv - u_0v_0 \\ &= (u_0 + \Delta u)(v_0 + \Delta v) - u_0v_0 \\ &= u_0v_u + u_0\Delta v + v_0\Delta u + \Delta u\Delta v - u_0v_0 \\ &= u_0\Delta v + v_0\Delta u + \Delta u\Delta v.\end{aligned}$$

[Notice how closely this resembles the infinitesimal version.] Now, we apply the definition[1] of the derivative as a limit:

$$\begin{aligned}\frac{d}{dx}(uv)\bigg|_{x=x_0} &= \lim_{\Delta x\to 0}\frac{\Delta(uv)}{\Delta x} \\ &= \lim_{\Delta x\to 0}\frac{u_0\Delta v + v_0\Delta u + \Delta u\Delta v}{\Delta x} \\ &= \lim_{\Delta x\to 0} u_0\frac{\Delta v}{\Delta x} + v_0\frac{\Delta u}{\Delta x} + \frac{\Delta u}{\Delta x}\cdot\frac{\Delta v}{\Delta x}\cdot\Delta x \\ &= u_0\,\frac{dv}{dx}\bigg|_{x=x_0} + v_0\,\frac{du}{dx}\bigg|_{x=x_0} + \left(\frac{du}{dx}\bigg|_{x=x_0}\right)\left(\frac{dv}{dx}\bigg|_{x=x_0}\right)\cdot 0 \\ &= u_0\,\frac{dv}{dx}\bigg|_{x=x_0} + v_0\,\frac{du}{dx}\bigg|_{x=x_0}.\end{aligned}$$ $\square$

Note the trick on the third line that was used to show that $\Delta u\Delta v/\Delta x \to 0$:

$$\frac{\Delta u\Delta v}{\Delta x} = \frac{\Delta u\Delta v\Delta x}{(\Delta x)^2} = \frac{\Delta u}{\Delta x}\cdot\frac{\Delta v}{\Delta x}\cdot\Delta x \to \frac{du}{dx}\cdot\frac{dv}{dx}\cdot 0 = 0$$

as $\Delta x \to 0$. This (sort of) gives a justification for the infinitesimal idea that $du\,dv = 0$.

4 The Chain Rule

Arguably the most important of all of these rules is the chain rule, which tells us how to take derivatives of compositions of functions. It states that if f and g are differentiable functions, then

$$(f\circ g)'(x) = f'(g(x))\cdot g'(x).$$

[1] Or rather, one of the equivalent definitions.

We can give a non-rigorous, infinitesimal derivation as follows: One (non-rigorous) definition of the derivative is that, if $y = f(x)$, then $f'(x)$ is the number such that

$$dy = f'(x)dx.$$

Now, suppose that $y = f(u)$ and $u = g(x)$, so that $y = f(u) = f(g(x))$. Then we have $dy = f'(u)du$ and $du = g'(x)dx$, so

$$\begin{aligned} dy &= f'(u)du \\ &= f'(u)g'(x)dx \\ &= f'(g(x))g'(x)dx. \end{aligned}$$

Hence,

$$\frac{dy}{dx} = f'(g(x))g'(x).$$

Example 3. Differentiate $(x+1)^{500}$.

Solution.

$$\begin{aligned} \frac{d}{dx}(x+1)^{500} &= 500(x+1)^{499} \cdot \frac{d}{dx}(x+1) \\ &= 500(x+1)^{499} \cdot 1 \\ &= 500(x+1)^{499}. \end{aligned}$$ □

It would have been possible, but very hard, to differentiate this by expanding out all 501 terms of the polynomial and then applying the techniques of the first section.

Assignment 16 (due Wednesday, 9 November)

Section 2.2, Problems 11–14 and 55–58. Follow the instructions. Remember, these problems (except for 57 and 58) are more about *how* you find the derivative than what derivative you find. The even-numbered problems will be graded carefully (although a certain amount of leeway will be provided on problem 58).

Section 2.3, Problems 11–14 and 23–26. (Hint: 23–26 are easier if you use the product rule.) The even-numbered problems will be graded carefully.

Translate the statement

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0)$$

into ε-δ language. (Hint: when you see $f'(x_0)$, treat it like ℓ. Also, treat $\Delta y/\Delta x$ as a function of Δx.) Then, use the resulting statement to prove the following:

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. if } |\Delta x| < \delta, \text{ then } \Delta y \text{ is within } \varepsilon|\Delta x| \text{ of } f'(x_0)\Delta x.$$

You will need to handle $\Delta x = 0$ as a separate case. This statement is a rigorous version of the statement that "When Δx is small, then Δy is approximately $\frac{dy}{dx}\Delta x$."

Assignment 17 (due Friday, 11 November)

From Section 2.3:

- Problems 5–8. Do each problem two ways—using the limit definition of your choice, and using the rules of differentiation (including the Chain Rule, if you find it helpful).
- Problems 17–20.
- Problems 31–32. Do not FOIL out the products; instead, use the product rule for differentiation.

The even-numbered problems will be graded carefully.

Section 2.5, Problems 1–4. Make sure it is clear, from your answer, how you are using the Chain Rule (see, for instance, Example 3 at the end of Lecture 18). Problems 2 and 4 will be graded carefully.

Give an ε-δ proof for each of the following:

1. Let f be the function defined by

$$f(x) = \begin{cases} 7x - 3 & \text{if } x \leq 0, \\ -\frac{1}{9}x - 3 & \text{if } x > 0. \end{cases}$$

   Show that $\lim_{x \to 0} f(x) = -3$.

2. Let f be the function defined by

$$f(x) = \begin{cases} -\frac{1}{7}x - \frac{18}{7} & \text{if } x < 3, \\ \frac{1}{6}x - \frac{7}{2} & \text{if } x \geq 3. \end{cases}$$

   Show that $\lim_{x \to 3} f(x) = -3$.

3. Let f be the function defined by

$$f(x) = \begin{cases} -\frac{1}{2}x + \frac{3}{2} & \text{if } x < -3, \\ 4 & \text{if } x = -3, \\ 3x + 12 & \text{if } x > -3. \end{cases}$$

   Show that $\lim_{x \to -3} f(x) = 3$.

Problems 1 and 3 will be graded carefully.

Suppose $y = f(x)$ and $f(x_0) = y_0$. A purely ε-δ version of the statement that "f is continuous at x_0" is given as follows:

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. if } |x - x_0| < \delta, \text{ then } |y - y_0| < \varepsilon.$$

Use this definition to prove the following fact:

> Suppose that
>
> $$\begin{aligned} u &= f(x), \\ u_0 &= f(x_0), \\ y &= g(u) = g(f(x)), \quad \text{and} \\ y_0 &= g(u_0) = g(f(x_0)). \end{aligned}$$
>
> If f is continuous at x_0 and g is continuous at u_0, then $g \circ f$ is continuous at x_0.

This will be graded carefully.

Part of Assignment 18

Assignment 18 will include giving ε-δ proofs of the following:

1. Let f be the function defined by

$$f(x) = \begin{cases} -8x + 10 & \text{if } x \leq 1, \\ 3x - 1 & \text{if } x > 1. \end{cases}$$

   Show that $\lim_{x \to 1} f(x) = 2$.

2. Let f be the function defined by

$$f(x) = \begin{cases} -4x+5 & \text{if } x < 1, \\ -\frac{1}{2}x + \frac{3}{2} & \text{if } x > 1. \end{cases}$$

Show that $\lim\limits_{x \to 1} f(x) = 1$.

3. Let f be the function defined by

$$f(x) = \begin{cases} -8x-2 & \text{if } x < 0, \\ -2 & \text{if } x = 0, \\ \frac{1}{7}x - 2 & \text{if } x > 0. \end{cases}$$

Show that $\lim\limits_{x \to 0} f(x) = -2$.

Test II around Wednesday, 16 November

Experienced teachers of Math 131 tell me that you will probably have a lot of papers and the like due around the time of the test, and consequently will not have a lot of time to study for it. Thus, I suggest you start studying now. You may also want to think in terms of "practicing" rather than "studying": redoing old quiz and homework problems (without looking at the solutions, if you have them, until afterwards) may be more helpful than simply reading over them.

Math 131, Lecture 19: The Chain Rule

Charles Staats

Wednesday, 9 November 2011

Important Note: There have been a few changes to Assignment 17. Don't use the version from Lecture 18.

1 Notes on the quiz

I've finally graded the quiz, and I wanted to make a few remarks.

1. Even apart from the one massive typo, the quiz directions were confusing. I apologize for this. Next time (tomorrow), I will make sure to take more time and care in designing the quiz.

   Because the instructions were confusing, the quiz was difficult to grade. Since the grades don't actually count for anything, I did not try too hard to assign grades based on what I thought was a "fair" assessment of your work. Instead, I mostly concentrated on assigning grades—and making comments—in whatever way I thought would be most useful to you.

2. In particular, on the delta-epsilon proof, assigning "fair" scores would have been virtually impossible since the statement of the problem was incorrect. Among other things, I deducted points for clear problems in the style of the δ-ε proofs.

3. If δs and εs should appear in your solution, I will say so explicitly in the problem instructions. In particular, if I ask you for the "limit-based definition of the derivative," I'm looking for something like

   $$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

   If I don't ask you for εs and δs, don't give them to me.

4. Do not—EVER—follow an expression like $\lim_{h\to 0}$ by an equals sign:

   $$\xcancel{f'(x) = \lim_{h\to 0} = \frac{f(x+h) - f(x)}{h}}$$

This simply makes no sense. The notation $\lim_{h\to 0}$ is supposed to represent the limit *of* an expression. If you instead follow it by an equals sign, it's like saying "The limit of is. . . ."

If you write this on a test, I will deduct points.

5. When I ask you to give an ε-δ proof for a limit of a piecewise-linear function, my secret goal is to get you to understand how you might go about proving the following statement:

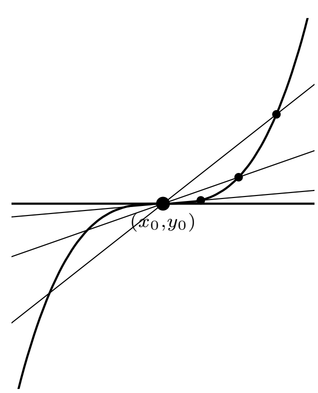
   > The two-sided limit exists and equals ℓ if and only if both the one-sided limits exist and equal ℓ.

   Some of you gave separate ε-δ proofs of both the one-sided limits, and then used the fact above to deduce the value of the two-sided limit. This is perfectly correct, and would have received full credit if I were grading the quiz for credit. However, since I was grading primarily to let you know whether you are prepared for this sort of problem on the test, I deducted a couple points. I very probably will give a problem like this on the test; if I do so, I will explicitly state that you are not allowed to use the statement above, so that I can justifiably deduct points if you do.

   Having reread the sentence above, I realized it sounds like I am looking for excuses to deduct points. This is NOT the case. When I give a particular problem, there are certain things I am trying to see if you understand. If you can do the problem correctly without understanding these things, then my whole purpose in giving the problem is compromised.

2 Computing the derivative from the definition

There are several limit-based definitions of the derivative. They all amount to saying "take the limit of the slope of the secant line between two points, as those two points get close together:"



The easiest definition to remember is probably

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

The easiest to compute with is probably based on the so-called "difference quotient":

$$\frac{dy}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

If I ask you to compute a derivative from the definition, I'm not mostly interested in whether you can find the answer. I may even phrase the question something like the following[1]:

> Use the definition of the derivative to prove that if f is the function defined by $f(x) = 1/x$, then $f'(x) = -1/x^2$.

As you may note, I am in fact giving you the "answer" here: $f'(x) = -1/x^2$. What I care about, when I ask a question like this, is whether you understand the definition well enough to use it.

Although we've done this example in Lecture 17, I'm going to repeat it here, since I will need it later in the lecture.

Example 1. Use the definition of the derivative to prove that if f is the function defined by $f(x) = \frac{1}{x}$, then $f'(x) = \frac{-1}{x^2}$.

Solution.

$$\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h \to 0} \frac{\frac{1}{x+h} - \frac{1}{x}}{h} \cdot \frac{x(x+h)}{x(x+h)} \\
&= \lim_{h \to 0} \frac{x - (x+h)}{hx(x+h)} \\
&= \lim_{h \to 0} \frac{-h}{hx(x+h)} \\
&= \lim_{h \to 0} \frac{-1}{x(x+h)} \\
&= \frac{-1}{x^2}.
\end{aligned}$$

□

If you want more examples of this sort of computation, you should review Lecture 17. (There's a version on Chalk that includes solutions to all of the examples.)

[1] NOTE: If I ask you this question on the test, the function f will be different.

3 Differentiating Quotients

We can use the Chain Rule together with the Product Rule and Example 1 (page 3) to differentiate quotients.

Example 2. Find $\frac{d}{dx}\left(\frac{1}{x-1}\right)$.

Recall, from Example 1, that $D_x(1/x) = -1/x^2$.

Solution.

$$\begin{aligned} \frac{d}{dx}\left(\frac{1}{x-1}\right) &= \frac{-1}{(x-1)^2} \cdot \frac{d}{dx}(x-1) \\ &= \frac{-1}{(x-1)^2}. \end{aligned}$$

□

4 Proof Sketch of the Chain Rule

Let y be a function of x. What does it mean to say that the derivative of y at x_0 is equal to a number m?

$$\begin{aligned} &\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = m \\ \iff &\forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.t. if } 0 < |\Delta x| < \delta, \text{ then } \left|\frac{\Delta y}{\Delta x} - m\right| < \varepsilon \\ \iff &\forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.t. if } 0 < |\Delta x| < \delta, \text{ then } \frac{|\Delta y - m\Delta x|}{|\Delta x|} < \varepsilon \\ \iff &\forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.t. if } 0 < |\Delta x| < \delta, \text{ then } |\Delta y - m\Delta x| < \varepsilon|\Delta x| \end{aligned}$$

Informally, this means that the statement $m = f'(x)$ is equivalent to the statement that

If the change in x is small, then $\Delta y \approx m\Delta x$.

In other words, near x_0, y is approximated by the tangent line. See Figure 1.

This suggests a (non-rigorous) definition of the derivative using infinitesimals: if $y = f(x)$, then $f'(x)$ is the number such that

$$dy = f'(x)\,dx.$$

This "definition" is based on the general notion that "if something is approximately true for small Δx, then it should be exactly true for dx because dx is so small." Thus, since $\Delta y \approx f'(x)\Delta x$, we get $dy = f'(x)\,dx$. This principle can get you in big trouble if applied indiscriminately, which is why using infinitesimals

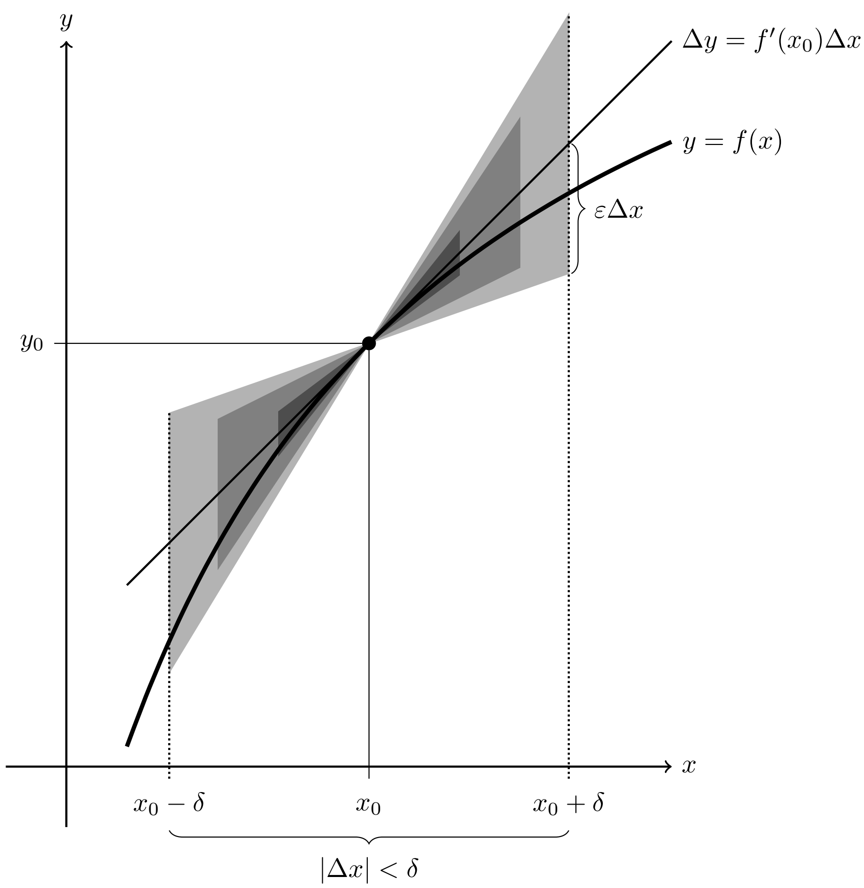


Figure 1: When Δx is small, then Δy is approximated by $f'(x_0)\Delta x$. In other words, for Δx small, the function is approximated by its tangent line (which is defined by $\Delta y = f'(x_0)\Delta x$). More precisely, the function is contained in a narrow cone about the tangent line. The width of the cone is controlled by ε. We can make the cone as narrow as we want ("arbitrarily narrow"), by making δ (and hence Δx) sufficiently small.

is "walking on clouds." But in many circumstances, it can give good intuition and correct results.

Now, suppose that $y = f(u)$ and $u = g(x)$, so that $y = f(u) = f(g(x))$. Then we have $dy = f'(u)\,du$ and $du = g'(x)\,dx$, so

$$\begin{aligned} dy &= f'(u)\,du \\ &= f'(u)g'(x)\,dx \\ &= f'(g(x))g'(x)\,dx. \end{aligned}$$

Hence, by the "infinitesimal definition of the derivative,"

$$\frac{dy}{dx} = f'(g(x))g'(x).$$

> Note: If I ask you on a test for the "Leibniz derivation of the Chain Rule" or the "Infinitesimal derivation of the Chain Rule," I am asking you, more or less, to give me the paragraph above.

Theorem. (Chain Rule) If f and g are differentiable functions, then $f \circ g$ is also differentiable, and

$$(f \circ g)'(x) = f'(g(x))g'(x).$$

The proof of the Chain Rule is to use εs and δs to say exactly what is meant by "approximately equal" in the argument

$$\begin{aligned} \Delta y &\approx f'(u)\Delta u \\ &\approx f'(u)g'(x)\Delta x \\ &= f'(g(x))g'(x)\Delta x. \end{aligned}$$

Unfortunately, there are two complications that have to be dealt with. The first is that, for technical reasons, we need an ε-δ definition for the derivative that allows $|\Delta x| = 0$. The following statement turns out to work:

$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.t. if } |\Delta x| < \delta, \text{ then } |\Delta y - f'(x_0)\Delta x| \leq \varepsilon |\Delta x|.$$

Comparing this to the earlier version, we got rid if the requirement $0 < |\Delta x|$ by changing the final $< \varepsilon|\Delta x|$ to $\leq \varepsilon|\Delta x|$. I don't want to explain why exactly we can do this, but anyone who has taken (and understood) an analysis course ought to be able to do it without much trouble.

The second complication is that the expression for δ in terms of ε turns out to be a bit ugly. For this reason, I will spare you the details. However, I hope I have convinced you that the basic idea of the proof of the Chain Rule is comprehensible, even if the technical details are a bit involved.

Assignment 17 (due Friday, 11 November)

From Section 2.3:

- Problems 5–8. Do each problem two ways—using the limit definition of your choice, and using the rules of differentiation (including the Chain Rule, if you find it helpful).
- Problems 17–20.
- Problems 31–32. Do not FOIL out the products; instead, use the product rule for differentiation.

The even-numbered problems will be graded carefully.

Section 2.5, Problems 1–4. Make sure it is clear, from your answer, how you are using the Chain Rule (see, for instance, Example 3 at the end of Lecture 18). Problems 2 and 4 will be graded carefully.

Give an ε-δ proof for each of the following. Do *not* use the fact that if both the one-sided limits exist and are equal, then the two-sided limit exists and is equal to both of them.

1. Let f be the function defined by
$$f(x) = \begin{cases} 7x - 3 & \text{if } x \leq 0, \\ -\frac{1}{9}x - 3 & \text{if } x > 0. \end{cases}$$
Show that $\lim_{x \to 0} f(x) = -3$.
2. Let f be the function defined by
$$f(x) = \begin{cases} -\frac{1}{7}x - \frac{18}{7} & \text{if } x < 3, \\ \frac{1}{6}x - \frac{7}{2} & \text{if } x \geq 3. \end{cases}$$
Show that $\lim_{x \to 3} f(x) = -3$.
3. Let f be the function defined by
$$f(x) = \begin{cases} -\frac{1}{2}x + \frac{3}{2} & \text{if } x < -3, \\ 4 & \text{if } x = -3, \\ 3x + 12 & \text{if } x > -3. \end{cases}$$
Show that $\lim_{x \to -3} f(x) = 3$.

Problems 1 and 3 will be graded carefully.

[NOTE: This is now a Bonus Problem.] Suppose $y = f(x)$ and $f(x_0) = y_0$. A purely ε-δ version of the statement that "f is continuous at x_0" is given as follows:

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. if } |x - x_0| < \delta, \text{ then } |y - y_0| < \varepsilon.$$

Use this definition to prove the following fact:

> Suppose that
>
> $$\begin{aligned} u &= f(x), \\ u_0 &= f(x_0), \\ y &= g(u) = g(f(x)), \quad \text{and} \\ y_0 &= g(u_0) = g(f(x_0)). \end{aligned}$$
>
> If f is continuous at x_0 and g is continuous at u_0, then $g \circ f$ is continuous at x_0.

Assignment 18 (due Monday, 14 November

Section 2.2, Problems 9, 10, 15, and 16. These problems are about the process of computing the derivative from the limit; finding the "answer" by another method will not receive full credit. You do not need to hand in any of these, but a similar problem will appear on the test.

Section 2.3, Problems 27–30. Use the Product Rule. Problems 28 and 30 will be graded carefully.

Section 2.5, Problems 5–8, 13–14, and 17–18. You do not need to show every single step, but it should be clear to the grader how you got to the answer. The even-numbered problems will be graded carefully.

Differentiate the following expressions with respect to x. (Hint: Apply the Chain Rule more than once.) You do not need to show every single step, but it should be clear to the grader how you got to the answer. You do not need to simplify the answer.

1. $\left(5(2x+1)^{361} - 17\right)^{42}$
2. $\left(1 - (1-2x)^{33}\right)^{1776}$

Both of these will be graded carefully. Since a similar problem may appear on the test, and this homework set will almost certainly not be graded before the test, you may want to ask to go over the answers in tutorial on Tuesday.

Give ε-δ proofs of the following facts. Do *not* use the fact that if both the one-sided limits exist and are equal, then the two-sided limit exists and is equal to both of them.

1. Let f be the function defined by

$$f(x) = \begin{cases} -8x + 10 & \text{if } x \leq 1, \\ 3x - 1 & \text{if } x > 1. \end{cases}$$

   Show that $\lim_{x \to 1} f(x) = 2$.

2. Let f be the function defined by

$$f(x) = \begin{cases} -4x + 5 & \text{if } x < 1, \\ -\frac{1}{2}x + \frac{3}{2} & \text{if } x > 1. \end{cases}$$

   Show that $\lim_{x \to 1} f(x) = 1$.

3. Let f be the function defined by

$$f(x) = \begin{cases} -8x - 2 & \text{if } x < 0, \\ -2 & \text{if } x = 0, \\ \frac{1}{7}x - 2 & \text{if } x > 0. \end{cases}$$

   Show that $\lim_{x \to 0} f(x) = -2$.

Problems 1 and 3 will be graded carefully.

Test II around Wednesday, 16 November

Experienced teachers of Math 131 tell me that you will probably have a lot of papers and the like due around the time of the test, and consequently will not have a lot of time to study for it. Thus, I suggest you start studying now. You may also want to think in terms of "practicing" rather than "studying": redoing old quiz and homework problems (without looking at the solutions, if you have them, until afterwards) may be more helpful than simply reading over them.

Math 131, Lecture 20: The Chain Rule, continued

Charles Staats

Friday, 11 November 2011

1 A couple notes on quizzes

I have a couple more notes inspired by the quizzes.

1.1 Concerning δ-ε proofs

First, concerning δ-ε proofs. There are a couple of commonsense rules that can alert you to when you are making a very bad choice of δ.

First, δ is always a positive number. If you ever find yourself writing something like $\delta = -\frac{1}{2}\varepsilon$, you should immediately realize you've done something wrong. The most way to get here was to forget the absolute value signs around the -2 when writing something like

$$|-2(x-1)| = |-2| \cdot |x-1|.$$

Second, δ is typically a *very small* positive number. If you write something like $\delta = 2 + \varepsilon$, then there is no possibility that δ will ever be less than 2. This is almost never what you want.

An error like demonstrates clear limits to your understanding of ε-δ proofs, and will probably cost you more points on a test than a less obvious error.

1.2 Infinitesimals are never equal to finite quantities

A finite quantity is something like $1+x$ without any d's in it. A ratio of infinitesimals like dy/dx is also a finite quantity. But if you multiply by a d(something) and don't divide by one, then your quantity is infinitesimal. You can equate two infinitesimals (e.g., $dy = f'(x)\,dx$) or two finite quantities (e.g., $dy/dx = f'(x)$). But if you ever have a finite quantity equal to an infinitesimal quantity, then you are doing something terribly wrong. In particular, for formulations of the product rule, we have

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx} \qquad \checkmark$$

$$d(uv) = u\,dv + v\,du \qquad \checkmark$$

$$\xcancel{\frac{d}{dx}(uv) = u\,dv + v\,du}$$

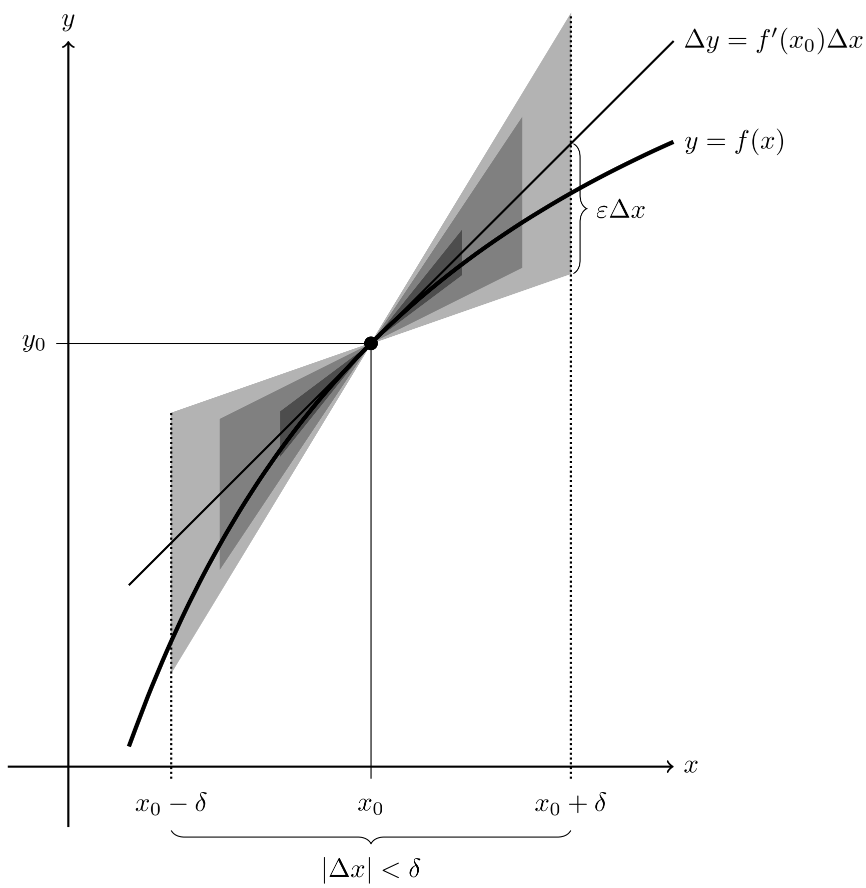


Figure 1: When Δx is small, then Δy is approximated by $f'(x_0)\Delta x$. In other words, for Δx small, the function is approximated by its tangent line (which is defined by $\Delta y = f'(x_0)\Delta x$). More precisely, the function is contained in a narrow cone about the tangent line. The width of the cone is controlled by ε. We can make the cone as narrow as we want ("arbitrarily narrow"), by making δ (and hence Δx) sufficiently small.

2 The Infinitesimal derivation of the Chain Rule

As you may recall from last lecture, the infinitesimal derivation of the Chain Rule goes something like this:

> Let $y = f(u)$ and $u = g(x)$. Then we have
>
> $$\begin{aligned} dy &= f'(u)\,\underbrace{du} \\ &= f'(\underbrace{\;u\;})\,\overbrace{g'(x)\,dx}, \qquad \text{since } du = g'(x)dx \\ &= f'\big(\overbrace{g(x)}\big)\,g'(x)\,dx \qquad \text{since } u = g(x). \end{aligned}$$
>
> Hence,
>
> $$\begin{aligned} \frac{dy}{dx} &= f'(g(x))g'(x), \qquad \text{i.e.,} \\ (f \circ g)'(x) &= f'(g(x))g'(x). \end{aligned}$$

The "infinitesimal statement" that $dy = f'(u)\,du$ corresponds to the "approximate statement" that $\Delta y \approx f'(u)\Delta u$. The basic idea behind the proof of the Chain Rule is to come up with a precise, ε-δ version of this "approximate statement," and then use that to turn the notion that

$$\Delta y \approx f'(u)\Delta u \approx f'(u)g'(x)\Delta x$$

into a precise proof. I will not repeat this ε-δ statement here, but I have included an illustration of it for your viewing pleasure in Figure 1.

3 Remembering the Chain Rule

Recall the Chain Rule, as stated in the last lecture:

Theorem. If f and g are differentiable functions, then

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

This is probably the form in which the Chain Rule is easiest to use, but it's kind of hard to remember. It becomes a lot easier to remember if we restate it in Leibniz notation. To do this, assume that $u = g(x)$ and $y = f(u) = f(g(x)) = (f \circ g)(x)$. Thus, we have

$$\begin{aligned} (f \circ g)'(x_0) &= \left.\frac{dy}{dx}\right|_{x=x_0} \\ f'(g(x_0) = f'(u_0) &= \left.\frac{dy}{du}\right|_{u=u_0} = \left.\frac{dy}{du}\right|_{u=g(x_0)} \\ g'(x_0) &= \left.\frac{du}{dx}\right|_{x=x_0} \end{aligned}$$

and the Chain Rule becomes

$$\left.\frac{dy}{dx}\right|_{x=x_0} = \left(\left.\frac{dy}{du}\right|_{u=g(x_0)}\right)\left(\left.\frac{du}{dx}\right|_{x=x_0}\right),$$

or more simply,

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

When written this way, the Chain rule seems completely obvious—just cancel the du's. This is *not* a great way to think about why the Chain Rule is actually true, because unlike most infinitesimal arguments, it cannot be turned into a rigorous proof. **If I ask you for the infinitesimal or Leibniz derivation of the Chain Rule on the test, the explanation here will *not* receive full credit.** However, it does make a good mnemonic device.

4 Using the Chain Rule

The textbook's section on the Chain Rule (Section 2.5) is actually not bad, and you might want to take a look at it (especially if you find my notes confusing). To quote the textbook, the key idea in applying the Chain Rule is that

> The last step in calculation corresponds to the first step in differentiation.

Example 1. Use the Chain Rule to differentiate $(2x+1)^3$.

Again quoting the textbook (more or less), the last step in the calculation is to cube something, so you start off by differentiating the cube function.

Solution (long version). Let

$$\begin{aligned} u &= 2x+1 \\ y &= u^3 = (2x+1)^3. \end{aligned}$$

Then

$$\begin{aligned} \frac{du}{dx} &= 2 \\ \frac{dy}{du} &= 3u^2 \end{aligned}$$

and so

$$\begin{aligned} \frac{dy}{dx} &= \frac{dy}{du}\frac{du}{dx} \\ &= 3u^2 \cdot 2 \\ &= 6(2x+1)^2, \end{aligned}$$

where the last step is obtained by substituting in $u = 2x+1$. □

Solution (short version).

$$\begin{aligned}\frac{d}{dx}(2x+1)^3 &= 3(2x+1)^2 \cdot \frac{d}{dx}(2x+1) \\ &= 3(2x+1)^2 \cdot 2 \\ &= 6(2x+1)^2.\end{aligned}$$

□

5 Differentiating Quotients

Recall that last lecture, we computed that

$$\frac{d}{dx}\frac{1}{x} = \frac{-1}{x^2}.$$

We can use this, together with the Chain Rule, to compute a lot of derivatives. To start with, we will take a look at x^n when n is a negative integer.

Example 2. Let $f(x) = x^{-m}$, where m is a positive integer. We may use the Chain Rule to compute $f'(x)$, as follows:

$$\begin{aligned}\frac{df}{dx} &= \frac{d}{dx}\left(\frac{1}{x^m}\right) \\ &= \frac{-1}{(x^m)^2} \cdot \frac{d}{dx}(x^m) \\ &= \frac{-1}{x^{2m}} \cdot mx^{m-1} \\ &= -m \cdot x^{-2m+(m-1)} \\ &= -m \cdot x^{-m-1}.\end{aligned}$$

If $n = -m$ is a negative integer, then we get

$$\frac{d}{dx}x^n = nx^{n-1}.$$

Thus, we have that the power rule holds for negative integers as well as positive integers. Since $D_x(x^0) = D_x(1) = 0 = 0 \cdot x^{-1}$ for $x \neq 0$, it might also be said that the power rule holds for 0.

Theorem. (Power Rule—all integers) If n is an integer, then

$$\frac{d}{dx}(x^n) = nx^{n-1}.$$

This should probably be memorized, but if you already have the one for positive integers memorized, that will probably not be difficult. (We will later show using implicit differentiation that the Power Rule holds whenever n is a rational number. It is in fact true even when n is irrational, although proving that requires logarithms.)

We can also use the Chain Rule, together with the Product Rule, to differentiate quotients.

Theorem. (Quotient Rule) Let f and g be differentiable functions. Then

$$\begin{aligned} D_x\left(\frac{f(x)}{g(x)}\right) &= \frac{g(x)D_x(f(x)) - f(x)D_x(g(x))}{g(x)^2} \\ &= \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}. \end{aligned}$$

You can either memorize the Quotient Rule, or remember how to differentiate quotients by combining the Product Rule with the Chain Rule. As long as you can differentiate quotients, I don't much care which method you use. If you do want to memorize this, the standard mnemonic is

> "Dee quotient equals bottom Dee top minus top Dee bottom, all over bottom squared."

However, if you use this mnemonic, remember not to equate infinitesimals with finite quantities. Either all the Dees should be D_x (derivative with respect to x, a finite quantity) or they should all be d (gives infinitesimals on both sides).

Proof.

$$\begin{aligned} D_x\left(\frac{f(x)}{g(x)}\right) &= D_x\left(f(x)\cdot\frac{1}{g(x)}\right) \\ &= f(x)D_x\left(\frac{1}{g(x)}\right) + \frac{1}{g(x)}D_x(f(x)) && \text{(product rule)} \\ &= f(x)\cdot\frac{-1}{(g(x))^2}\cdot D_x(g(x)) + \frac{f'(x)}{g(x)} && \text{(chain rule)} \\ &= \frac{-f(x)g'(x)}{g(x)^2} + \frac{f'(x)}{g(x)} \\ &= \frac{-f(x)g'(x) + f'(x)g(x)}{g(x)^2} \\ &= \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}. \end{aligned}$$

□

Example 3. Let

$$f(x) = \frac{x+1}{x-1}.$$

Compute $f'(x)$.

Without the quotient rule.

$$\begin{aligned}
\frac{d}{dx}\left(\frac{x+1}{x-1}\right) &= \frac{d}{dx}\left((x+1)\cdot\frac{1}{x-1}\right) \\
&= (x+1)\frac{d}{dx}\left(\frac{1}{x-1}\right) \\
&\quad + \frac{1}{x-1}\frac{d}{dx}(x+1) && \text{(Product Rule)} \\
&= (x+1)\cdot\frac{-1}{(x-1)^2}\cdot\frac{d}{dx}(x-1) \\
&\quad + \frac{1}{x-1}\cdot 1 && \text{(Chain Rule)} \\
&= -\frac{x+1}{(x-1)^2} + \frac{1}{x-1} \\
&= \frac{-(x+1)+(x-1)}{(x-1)^2} \\
&= \frac{-2}{(x-1)^2}.
\end{aligned}$$

□

With the quotient rule.

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

□

Assignment 18 (due Monday, 14 November)

Section 2.2, Problems 9, 10, 15, and 16. These problems are about the process of computing the derivative from the limit; finding the "answer" by another method will not receive full credit. You do not need to hand in any of these, but a similar problem will appear on the test.

Section 2.3, Problems 27–30. Use the Product Rule. Problems 28 and 30 will be graded carefully.

Section 2.5, Problems 5–8, 13–14, and 17–18. You do not need to show every single step, but it should be clear to the grader how you got to the answer. The even-numbered problems will be graded carefully.

Differentiate the following expressions with respect to x. (Hint: Apply the Chain Rule more than once.) You do not need to show every single step, but it should be clear to the grader how you got to the answer. You do not need to simplify the answer.

1. $\left(5(2x+1)^{361} - 17\right)^{42}$
2. $\left(1 - (1-2x)^{33}\right)^{1776}$

Both of these will be graded carefully. Since a similar problem may appear on the test, and this homework set will almost certainly not be graded before the test, you may want to ask to go over the answers in tutorial on Tuesday.

Give ε-δ proofs of the following facts. Do *not* use the fact that if both the one-sided limits exist and are equal, then the two-sided limit exists and is equal to both of them.

1. Let f be the function defined by
$$f(x) = \begin{cases} -8x+10 & \text{if } x \leq 1, \\ 3x-1 & \text{if } x > 1. \end{cases}$$
Show that $\lim_{x \to 1} f(x) = 2$.
2. Let f be the function defined by
$$f(x) = \begin{cases} -4x+5 & \text{if } x < 1, \\ -\frac{1}{2}x + \frac{3}{2} & \text{if } x > 1. \end{cases}$$
Show that $\lim_{x \to 1} f(x) = 1$.

3. Let f be the function defined by

$$f(x) = \begin{cases} -8x - 2 & \text{if } x < 0, \\ -2 & \text{if } x = 0, \\ \frac{1}{7}x - 2 & \text{if } x > 0. \end{cases}$$

   Show that $\lim\limits_{x \to 0} f(x) = -2$.

Problems 1 and 3 will be graded carefully.

Test II around Wednesday, 16 November

Experienced teachers of Math 131 tell me that you will probably have a lot of papers and the like due around the time of the test, and consequently will not have a lot of time to study for it. Thus, I suggest you start studying now. You may also want to think in terms of "practicing" rather than "studying": redoing old quiz and homework problems (without looking at the solutions, if you have them, until afterwards) may be more helpful than simply reading over them.

Assignment 19 (due Friday, 18 November)

Section 2.6, Problems 1–4. Problems 2 and 4 will be graded carefully.

Assume that f is a differentiable function. Consider the two functions g and h defined by

$$g(t) = 2f(t),$$
$$h(t) = f(2t).$$

You may want to check your answers below by considering the specific cases of $f(t) = t$ and $f(t) = t^2$.

1. Explain how to obtain the graphs of g and h from the graph of f by shrinking/stretching.

2. Compute g' and h' in terms of f'. (Hint: they are NOT the same.)

3. Explain how to obtain the graphs of g' and h' from the graph of f' by shrinking/stretching.

All three of these will all be graded carefully.

Math 131, Lecture 21

Charles Staats

Monday, 14 November 2011

1 Differentiability and Continuity

There's a theoretical point that I've sort of hand-waved over up to now, but that probably needs to be addressed. If you recall, the definition of the derivative (or at least, one of the definitions) is

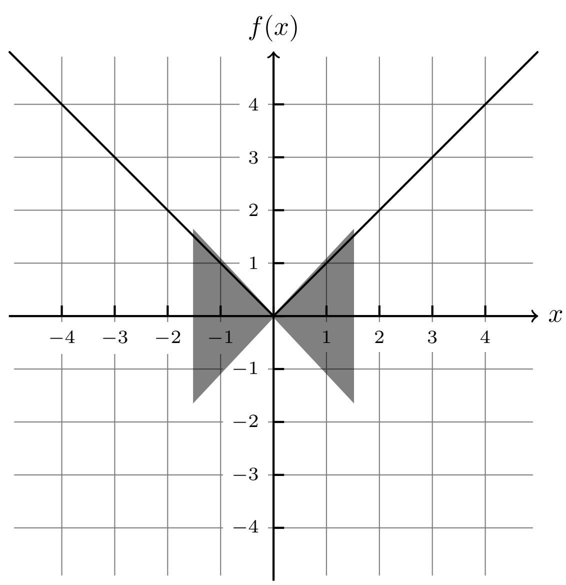
$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

However, an important point about limits is that they don't always exist. Similarly, derivatives do not always exist.

Example 1. Let f be the absolute value function; i.e., f is defined by

$$f(x) = |x| = \begin{cases} -x & \text{if } x < 0, \\ x & \text{if } x \geq 0. \end{cases}$$

Then $f'(0)$ does not exist. Geometrically, we can see this because it is not possible to draw a narrow cone centered on $(0, 0)$ that contains the graph of f:



More formally, we see that

$$\lim_{h\to 0^-} \frac{|0+h|-|0|}{h} = \lim_{h\to 0} \frac{-h}{h} = -1$$
$$\lim_{h\to 0^+} \frac{|0+h|-|0|}{h} = \lim_{h\to 0} \frac{h}{h} = 1.$$

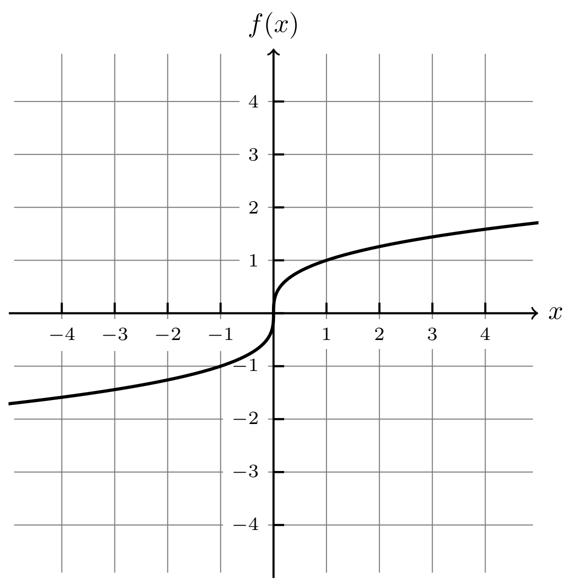
Since the one-sided limits are not equal, the two-sided limit

$$\lim_{h\to 0} \frac{|0+h|-|0|}{h} = f'(0)$$

does not exist.

Definition. We say that a function f is *differentiable at* x_0 if f is defined at x_0 and the derivative $f'(x_0)$ exists (and is finite). We say that f is *differentiable* if it is differentiable at every point of its domain.

Example 2. Consider the function f defined by $f(x) = \sqrt[3]{x}$:



You should not find it hard to believe that the tangent line to f at the origin is vertical—i.e., a line with slope infinity. Correspondingly, if one evaluates

$$f'(0) = \lim_{h\to 0} \frac{(0+h)^{1/3} - 0^{1/3}}{h},$$

one will find that the limit is ∞. Since this derivative is not *finite*, we still say that f is not differentiable at 0.

One reason for this convention is that the Chain Rule does not work here: if it did, it would tell us that the derivative of $\sqrt[3]{g(x)}$ at $x = 0$ is $\infty \cdot g'(0)$. Say that

$g(x) = x^3$; then we know that $\sqrt[3]{g(x)} = x$, so the derivative at 0 should be 1; but the chain rule would tell us that this derivative is $\infty \cdot 0$, which does not make sense. However, because the Chain Rule only applies when both functions are differentiable, and $\sqrt[3]{}$ is not differentiable, we don't run into a contradiction.

As I have said before, the "main point" of functions is, more or less, that they give us a way to talk about things that we don't have formulas for. Thus, if we have a problem, we might be able to show that there is a function that gives its solution, even if there is no formula for the solution. Once we've shown that the solution is given by a function, we can ask how "nice" the function is: Is it continuous? Is it differentiable? In this situation, we would probably find the following theorem very interesting:

Theorem. Let f be a function. If f is differentiable at x_0, then f is continuous at x_0.

Proof. Assume f is differentiable at x_0. Then f is defined at x_0, by definition of differentiability.

Let $y = f(x)$, $y_0 = f(x_0)$, and note that

$$\begin{aligned} y &= y_0 + \Delta y \\ &= y_0 + \frac{\Delta y}{\Delta x} \cdot \Delta x. \end{aligned}$$

Hence,

$$\begin{aligned} \lim_{x \to x_0} f(x) &= \lim_{\Delta x \to 0} y \\ &= \lim_{\Delta x \to 0} y_0 + \frac{\Delta y}{\Delta x} \cdot \Delta x \\ &= y_0 + \left(\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \right) \cdot \left(\lim_{\Delta x \to 0} \Delta x \right) \\ &= y_0 + \left(\left. \frac{dy}{dx} \right|_{x = x_0} \right) \cdot 0 \\ &= y_0 = f(x_0). \end{aligned}$$

□

Note: In the proof above, if the $dy/dx|_{x=x_0}$ did not exist (as a finite number), then the limit

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

would not have made sense, and so the Main Limit Theorem would not have been applicable.

For our purposes in this class, the most important use of this theorem may be a way to tell when a function is *not* differentiable. For this, we use the contrapositive:

> If f is discontinuous at x_0, then f is not differentiable at x_0.

Warning. *It is quite possible for a function to be continuous but not differentiable. For instance, our earlier examples $f(x) = |x|$ and $g(x) = \sqrt[3]{x}$ are both continuous, but neither is differentiable at* 0*.*

2 Higher derivatives

Given a differentiable function f, its derivative f' is also a function. This function f' may itself be differentiable, and have a derivative of its own, which we call f''—the *second derivative*. The derivative of f'', if it exists, is denoted f''', and called the *third derivative of* f.

Example 3. Let f be the function defined by $f(x) = x^5$. Find the first, second, and third derivatives of f.

Solution.

$$\begin{aligned} f'(x) &= 5x^4 \\ f''(x) &= 5 \cdot 4x^3 = 20x^3 \\ f'''(x) &= 20 \cdot 3x^2 = 60x^2. \end{aligned}$$

□

We can, of course, proceed to take higher derivatives than just the third derivative. But since something like $f'''''''(x)$ would be rather hard to read, we denote, e.g., the seventh derivative of f by $f^{(7)}$. There are several other notations:

read aloud	prime notation	D notation	Leibniz notation
the n^{th} derivative of f with respect to x	$f^{(n)}(x)$	$D_x^n(f(x))$	$\dfrac{d^n f}{dx^n}$

The Leibniz notation is based on the idea that $\dfrac{d}{dx}\left(\dfrac{dy}{dx}\right)$ should be written as $\dfrac{d^2y}{dx^2}$. Unlike most other versions of the Leibniz notation, this is purely a mnemonic device; trying to think about this as the "quotient" of "infinitesimal" quantities d^2y and dx^2 ends up just giving a mess.

Test Wednesday, 16 November

Test II will be on Wednesday. It will cover the lectures up through Lecture 20 (Friday), and the homework up through Assignment 18 (due today). You should also look at the quizzes (graded and ungraded) up to and including tomorrow's quiz. Although there will be one problem involving limits and ε-δ proofs, the emphasis of the test will be on derivatives, rather than limits. (Note, however, that you will be asked to find a derivative from the definition, which will require you to evaluate a limit.)

If you can do all the quiz problems without consulting anyone or anything, you should do well. When you've mastered the quiz problems, there still may be some additional benefit from practicing homework problems. Bonus problems will not be tested.

Last Tuesday, I sent out an e-mail outlining what sorts of problems I was thinking about putting on the test. As promised, that list is not perfect—some of the problems their did not make it, and one or two things that ended up on the test were not on the list. However, it's still a fairly good study guide.

The lecture notes, in addition to trying to explain the material, discuss some pitfalls that could cost you points, even if you think you know what you are doing.

On the last test, a lot of people who got basically the right answers still lost points, because what they wrote did not show me that they truly understood what was going on. If you want to know exactly how much you need to write to get credit on certain kinds of problems, please come to office hours (Tuesday, 3–4pm) and ask me. If you cannot make my office hours, I'll be happy to make an appointment with you.

Assignment 19 (due Friday, 18 November)

Section 2.6, Problems 1–4. Problems 2 and 4 will be graded carefully.

Assume that f is a differentiable function. Consider the two functions g and h defined by

$$g(t) = 2f(t),$$
$$h(t) = f(2t).$$

You may want to check your answers below by considering the specific cases of $f(t) = t$ and $f(t) = t^2$.

1. Explain how to obtain the graphs of g and h from the graph of f by shrinking/stretching.
2. Compute g' and h' in terms of f'. (Hint: they are NOT the same.)
3. Explain how to obtain the graphs of g' and h' from the graph of f' by shrinking/stretching.

All three of these will all be graded carefully.

Math 131, Lecture 22

Charles Staats

Friday, 18 November 2011

Organizational matters

- Tutorial scheduling for next quarter
- Vote: Do you want an assignment due Wednesday, November 23 (the day before Thanksgiving), or a double assignment due Monday, November 28?

1 Higher derivatives

Given a differentiable function f, its derivative f' is also a function. This function f' may itself be differentiable, and have a derivative of its own, which we call f''—the *second derivative*. The derivative of f'', if it exists, is denoted f''', and called the *third derivative of f*.

Example 1. Let f be the function defined by $f(x) = x^5$. Find the first, second, and third derivatives of f.

Solution.

$$\begin{aligned} f'(x) &= 5x^4 \\ f''(x) &= 5 \cdot 4x^3 = 20x^3 \\ f'''(x) &= 20 \cdot 3x^2 = 60x^2. \end{aligned}$$

□

We can, of course, proceed to take higher derivatives than just the third derivative. But since something like $f'''''''(x)$ would be rather hard to read, we denote, e.g., the seventh derivative of f by $f^{(7)}$. There are several other notations:

read aloud	prime notation	D notation	Leibniz notation
the n^{th} derivative of f with respect to x	$f^{(n)}(x)$	$D_x^n(f(x))$	$\dfrac{d^n f}{dx^n}$

The Leibniz notation is based on the idea that $\frac{d}{dx}\left(\frac{dy}{dx}\right)$ should be written as $\frac{d^2y}{dx^2}$. Unlike most other versions of the Leibniz notation, this is purely a mnemonic device; trying to think about this as the "quotient" of "infinitesimal" quantities d^2y and dx^2 ends up just giving a mess.

2 Acceleration

If you recall from when we first introduced the derivative, the first motivation I gave was that "the derivative is the rate of change of position with respect to time." If x represents position, then

$$\frac{dx}{dt}$$

represents velocity. One of the most important instances of a higher derivative is *acceleration*, or the rate of change of velocity with respect to time:

$$\frac{d^2x}{dt^2}.$$

If the velocity of an object is increasing, then its acceleration is positive; if the velocity is decreasing, then the acceleration is negative.

Intuitively, we are inclined to think that something is "accelerating" if it is "getting faster," and "decelerating" if it is "slowing down." This intuition can be useful, but it is also dangerous. If an object has positive velocity (i.e., moving to the right), but negative acceleration, then its velocity will decrease to zero, and continue to decrease to be negative; i.e., the object will start moving to the left. We could say that the object decelerates to a stop, and then accelerates in the opposite direction; however, this is deceptive, because the (negative) acceleration is exactly the same before, during, and after the instant at which the object is "stopped."

For another example, consider what happens when an object is tossed upwards. We might be inclined to say that under the force of gravity, it decelerates until it reaches the apex of its path, and then starts falling downward. But really, the acceleration is the same (negative) from the moment the object leaves the hand. Thus, it actually makes more sense to say that the object is falling from the instant it leaves the hand—even while it is still moving upward (i.e., has positive velocity).

One time (in middle school, I think), I was in an auditorium with a bunch of other students listening to an astronaut speak. At one point, he asked us why, when an astronaut in a spaceship "drops" something, it floats rather than falling. The auditorium shook as everyone in the audience shouted, "No gravity!" The astronaut replied, "Everyone who just said 'no gravity' is 100% wrong." If there were no gravity, then the spaceship would not be orbiting the earth; instead, it would be traveling away from the earth in a straight line, never to return. The

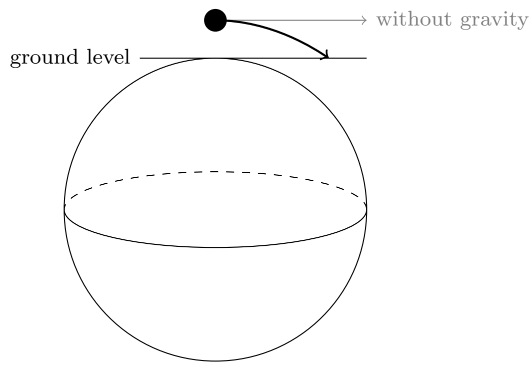


Figure 1: The spaceship is falling, but it's moving sidewise so quickly that by the time it reaches "ground level," it has not actually gotten any closer to the surface of the earth.

reason, he said, that an object dropped inside the spaceship appears to float is that the object, the astronaut, and the entire spaceship are already falling. The only reason the spaceship does not reach the ground is that by the time it reaches "ground level," it's moved so far horizontally that the ground has dropped out from beneath it.

For trajectories short enough that we can pretend the earth is flat, the general rule is the following:

Law of Falling Bodies. If an object is under no influences[1] but that of gravity, then its vertical acceleration is a constant $g \approx -10\frac{m}{s^2}$.

If the acceleration g is measured in feet per second squared rather than meters per second squared, its value is approximately -32. In either case, this acceleration is negative, because the object's velocity is decreasing. (If the object is moving downward, then its velocity is already negative, and is becoming more negative.)

I've called this the Law of Falling Bodies rather than the Law of Gravity because gravity generally refers to a deeper phenomenon discovered by Isaac Newton, whereas a version of the Law of Falling Bodies was known earlier to Galileo. (Who was forced to deal with average acceleration, because he did not know about derivatives.)

If we translate the Law of Falling Bodies into mathematical notation, we obtain the equation

$$\frac{d^2y}{dt^2} = -10,$$

where y is the vertical position of the object. This is a very simple example of what is called a *differential equation*; to "solve" the differential equation, we

[1] If this were a physics course, we'd use the word "forces."

what functions $y(t)$ would make it true. It is not hard to verify that for any choice of a and b, the function

$$y(t) = -5t^2 + at + b$$

is a solution to the differential equation:

$$\begin{aligned} y'(t) &= -10t + a \\ y''(t) &= -10. \end{aligned}$$

As it turns out, these are the only functions that satisfy this differential equation, although we will not see why until next quarter. Thus, any time you toss or drop an object, its vertical position is described by

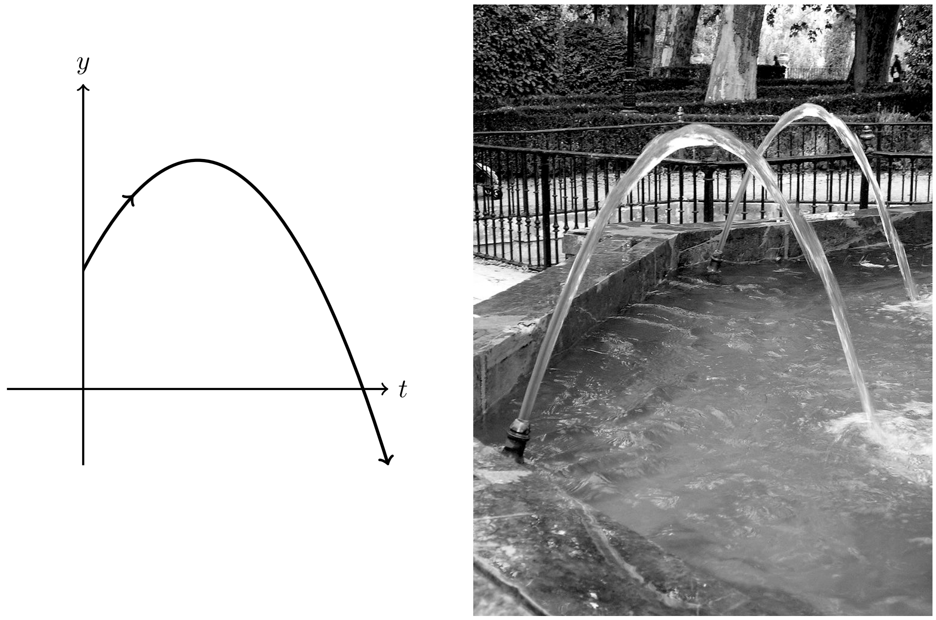
$$y = -5t^2 + at + b$$

for some choice of a and b. Note that there are many different paths possible, since there are many different values of a and b. This is good, since there are many different paths falling bodies can follow in real life. (If you throw a piece of chalk up, it will follow a different path from the piece of chalk you throw down, but both paths can be described by the equation $y = -5t^2 + at + b$, for some (different) values of a and b.)

However, no matter what a and b are, $y = -5t^2 + at + b$ is always some sort of upside-down parabola. Correspondingly, a falling object always moves in some form of upside-down parabola; see Figure 2a. If you imagine that your object is a droplet of water, and you string a bunch of these "objects" together in a continuous stream, you can see the whole path at once, as in Figure 2b. Notice how much more interesting nature's parabola is than the stark, abstract curve given in 2a. The water's arc seem to scintillate with reflected light; cords of water seem to twist together, like the muscles in a Michelangelo drawing of an arm.

In the textbook, it essentially just gives you the equation for the position, say $y = -5t^2 + t + 1$, and asks you to calculate the acceleration. And as it turns out, the acceleration is constantly -10 (or perhaps -32, since the textbook seems to like feet more than meters). While finding the acceleration from the position function is a perfectly good exercise, it somehow feels backwards. In some sense, the basic statement is that the vertical acceleration of a falling object is constantly -10; this basic fact is the cause of the effect that the object travels in a parabola given by $y = -5t^2 + at + b$. By starting off with the path and deducing the acceleration, it feels as though you are mixing up the cause and the effect.

One final note: If you look more closely at Figure 2a, you will see that the horizontal axis is indicating *time*. On the other hand, in Figure 2b, the "horizontal axis," such as it is, clearly is given by horizontal *position*, or distance (more or less). Since the graph does not actually tell you where the object is horizontally at a given time, it is not entirely clear why the "parabola" description should be accurate; the graph could just as easily describe an object that



(a) An object with constant negative acceleration moves in an upside-down parabola.

(b) Parabolic trajectory of water. By GuidoB. Modified (primarily to make it grayscale). This image is licensed under a Creative Commons Attribution-Share Alike 3.0 Unported license; see `http://creativecommons.org/licenses/by-sa/3.0/deed.en`.

Figure 2: Parabolas in theory and in practice

goes straight up and straight back down with no “sideways” movement. For the moment, I'm just going to ignore this discrepancy. We may, or may not, discuss it when discussing related rates.

3 Implicit Differentiation

Suppose we know, or suspect, that y is a differentiable function of x. We don't have a formula for y, but we may know that y and x satisfy some relation, for instance,

$$y^2 + x^2 = 1.$$

Often, we can use this, together with the chain rule, to figure out what the derivative of y must be (assuming it has one). In the example at hand, we differentiate both sides with respect to x, and then solve for the derivative $D_x y$:

$$\begin{aligned} D_x(y^2 + x^2) &= D_x(1) \\ 2y \cdot D_x(y) + 2x &= 0 \\ 2y D_x y &= -2x \\ D_x y &= \frac{-2x}{2y} = \frac{-x}{y}. \end{aligned}$$

This expression for the derivative $D_x y$ has a y in it as well as an x, which, as the book says, can be "a nuisance." However, it can nevertheless be quite useful. If we should happen to know that the value of y at a point x_0 is y_0, then we can use this to calculate $D_x y = dy/dx$ at the point (x_0, y_0), assuming this derivative exists.

Assignment 20 (due Monday, 21 November)

Section 2.3, Problems 39–40. Problem 40 will be graded carefully.

Section 2.6, Problems 7–10. Problems 8 and 10 will be graded carefully.

Section 2.7, Problems 1–2. Your expression for $D_x y$ may include both xs and ys. Problem 2 will be graded carefully.

The following argument purports to show that every function is continuous:

> Let f be a function, and x_0 any point in its domain. We show that f is continuous at x_0.
>
> $$\begin{aligned}\lim_{x\to x_0} f(x) &= \lim_{x\to x_0} f(x_0) + [f(x) - f(x_0)] \\ &= \lim_{x\to x_0} f(x_0) + \frac{f(x)-f(x_0)}{x-x_0}\cdot(x-x_0) \\ &= f(x_0) + \left(\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}\right)\cdot\left(\lim_{x\to x_0} x - x_0\right) \\ &= f(x_0) + \left(\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}\right)\cdot 0 \\ &= f(x_0) + 0,\end{aligned}$$
>
> since anything times zero is zero. Thus, f is continuous at x_0. Moreover, since the same argument applies to every point x_0 in the domain of f, we know that f is continuous at every point in its domain. In other words, f is continuous.

On the other hand, we know that not every function is continuous. Thus, there must be a flaw in the argument. What is it? (Hint: this argument *can* be used to show that every *differentiable* function is continuous.)

Part(?) of Assignment 21 (due Wednesday, 23 November)

[Note: If you won't be in class on the day before Thanksgiving, then some time before class, put the homework in my mailbox (in the Eckhart basement). Also, send me an email so that I know you have done this.]

Section 2.6, Problems 11–12 and 20–21. Problems 12 and 21 will be graded carefully.

Section 2.7, Problems 3–6 and 19–20. The even-numbered problems will be graded carefully.

("Semi-bonus problem") Suppose that

$$\lim_{x \to c^-} f(x) = \ell = \lim_{x \to c^+} f(x);$$

i.e., both the one-sided limits are defined, and they are equal. Use the ε-δ definition of the limit to show that the two-sided limit is also defined and equal to ℓ, i.e., that

$$\lim_{x \to c} f(x) = \ell.$$

The technique involved should be similar to that used to give ε-δ proofs for piecewise linear functions.

This is a "semi-bonus problem" in the following sense:

- If you do not seriously attempt it, you will not receive full credit on their homework.
- If you seriously attempt it, you will receive full credit for it (although your actual homework grade will, of course, depend on the other homework problems).
- If you get it right, you will receive a bonus point on the homework.

Math 131, Lecture 23

Charles Staats

Monday, 21 November 2011

1 Implicit Differentiation: How to differentiate a function we don't know

All of the "exercises" for differentiation so far have been based on differentiating *formulas*. However, many of the *rules* for differentiation (most especially, the chain rule) are much more general than this: they deal with differentiating *functions*. And, as you may recall, kind of the whole point of functions is that they are not necessarily given by formulas. We have not really explored this very far, because most of the functions we could talk about were, in fact, given by formulas. But there is another way: we can define a function as a *solution* to something. For instance, the $\sqrt{\ }$ function is really defined by

$$\sqrt{x} = \text{the nonnegative number } y \text{ such that } y^2 = x.$$

In other words, $\sqrt{x}$ is just a fancy way of writing "the (nonnegative) solution to the equation $y^2 = x$." And we can go back to this basic definition to differentiate the square root function:

Example 1. Suppose $y = \sqrt{x}$. Find an expression for dy/dx.

Solution. We assume, first of all, that $\sqrt{x}$ is in fact differentiable; without this assumption, there is not much we can do. We then go back to the basic equation that $y^2 = x$ and apply the Chain Rule:

$$\begin{aligned} y^2 &= x \\ \frac{d}{dx}y^2 &= \frac{d}{dx}x \\ 2y \cdot \frac{dy}{dx} &= 1 \\ \frac{dy}{dx} &= \frac{1}{2y} = \frac{1}{2\sqrt{x}}. \end{aligned}$$

Thus, assuming that the function f taking $x \mapsto \sqrt{x}$ is differentiable, its derivative f' is necessarily given by

$$f'(x) = \frac{1}{2\sqrt{x}}.$$

□

Exercise 2. Use implicit differentiation to show that if $y = -\sqrt{x}$, then

$$\frac{dy}{dx} = -\frac{1}{2\sqrt{x}}.$$

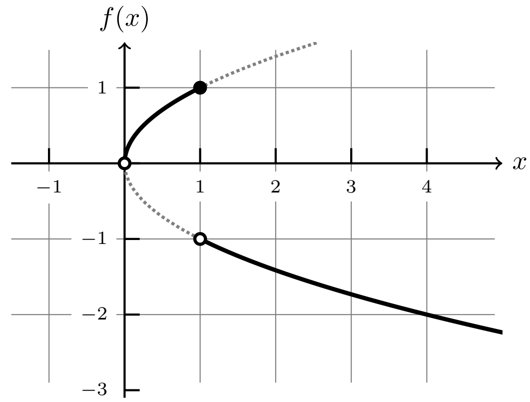
Solution.

Note that there is a problem we never dealt with here: we never actually showed that $f(x) = \sqrt{x}$ is differentiable. We only figured out what its derivative must be, *if* the derivative exists. There are a couple ways to solve this problem.

- It is possible to compute the derivative of $\sqrt{x}$ directly from the definition of the derivative (i.e., as the limit of the difference quotient). This is done earlier in the textbook. However, this is a way to avoid using implicit differentiation; what we really want is a way to show that implicit differentiation works.
- There is a theorem called the "Implicit Function Theorem" that states, roughly, that if implicit differentiation gives a reasonable answer, then the equation in question does in fact have a solution $y = f(x)$ where f is a differentiable function. This is kind of like the Main Limit Theorem: If the process gives a reasonable answer, then we know that must be the right answer; but if the process does not give a reasonable answer, we don't know anything.

The Implicit Function Theorem may seem to be the answer to our problems, but there are subtleties even here. First, the actual statement of the theorem is something that I find confusing, so I very much doubt that you want to see it. Second, while the Implicit Function Theorem can guarantee that *some* solutions are differentiable (in this case, $f(x) = \sqrt{x}$ and $f(x) = -\sqrt{x}$ are both solutions to $f(x)^2 = x$ that are differentiable for $x > 0$), there will also be other solutions that are *not* differentiable. For instance, if f is the function defined by

$$f(x) = \begin{cases} \sqrt{x} & \text{if } 0 < x \leq 1, \\ -\sqrt{x} & \text{if } x > 1, \end{cases}$$



then $y = f(x)$ is also a solution to the equation $y^2 = x$ for all $x > 0$, but f is not even continuous, much less differentiable. We will not try to explain why the Implicit Function Theorem applies for some "solutions," but not to others. Instead, we will adopt a "third way":

- Ignore the difficulties and just assume implicit differentiation works. Any function we encounter "naturally" in this course[1] is going to work out just fine.

In essence, we've reached a point where the skyscraper just gets too convoluted to deal with, so we're going to continue walking on clouds.

There's one more very important result we want to obtain using implicit differentiation. Recall that we proved the Power Rule, $D_x(x^n) = nx^{n-1}$, whenever n is an integer. We're now going to that this holds, not just for integers, but for rational numbers.

Theorem. (Power Rule for rational exponents) Let r be any rational number. Then

$$D_x(x^r) = rx^{r-1}.$$

Incomplete Proof. Since r is a rational number (i.e., a "ratio" of two integers), we may write

$$r = \frac{p}{q},$$

for some integers p, q, where $q \neq 0$. By definition, $y = x^{p/q}$ is a solution to the equation

$$y^q = x^p.$$

[1]That is, any function that has not been explicitly designed to cause problems.

Applying implicit differentiation, together with the power rule for integer exponents, we see that

$$\begin{aligned}
qy^{q-1}\frac{dy}{dx} &= px^{p-1} \\
\frac{dy}{dx} &= \frac{px^{p-1}}{qy^{q-1}} \\
&= \frac{p}{q}\cdot\frac{x^{p-1}}{(x^r)^{q-1}} \\
&= r\cdot x^{(p-1)-r(q-1)} \\
&= r\cdot x^{p-1-(p/q)(q-1)} \\
&= r\cdot x^{p-1-p+p/q} \\
&= r\cdot x^{-1+p/q} \\
&= r\cdot x^{r-1}.
\end{aligned}$$

□

The key point of this proof is that we could apply the power rule to x^p and y^q, because we already knew the power rule for integer exponents, and p, q are integers. This proof is incomplete in that we have not really turned implicit differentiation into a rigorous technique, so we can't use it in "real" proofs.

I commented at one point that calculus is "supposed" to work exactly the same for rational and irrational numbers. Thus, it seems peculiar that we have a rule that only seems to work for rational numbers. In fact, as it turns out, the Power Rule does hold for all real exponents—rational or irrational. There's even a nice, elegant proof that does not care whether r is rational or irrational. Unfortunately, this proof uses logarithms, so we won't see it for some time (if at all). Thus, for now, all our powers will be rational.

2 Some potential pitfalls: numbers, functions, and expressions

When I first introduced functions, I made a big deal of the fact that f is a function, but $f(x)$ is just a number (albeit one that we do not yet know). In terms of this distinction, differentiation is something we do to *functions*, not *numbers*. Thus, Df, the "derivative of f," is a function, but $Df(x)$ would be the "derivative of a number," which does not make any sense. Unfortunately, this distinction has become somewhat blurred when we write things like

$$\frac{d}{dx}(x^2+1).$$

What we really mean here is "the derivative of the function that maps $x \mapsto x^2+1$." The x in d/dx tells us that x is just a "dummy variable," and so the input is really just a function. When we write the answer as $2x$, it is even harder to tell that we mean “the function mapping $x \mapsto 2x$” rather than simply “the number $2x$.”

So far, this section has been entirely theoretical, but there is a practical, computational issue as well. Suppose someone asks you to calculate the derivative of $x^2 + 1$ at $x = 2$. You may be tempted to substitute in $x = 2$ *before* differentiating, which would be a disaster. You'd be differentiating a number rather than a function; you'd probably try to treat it as the constant function $2^2 + 1 = 5$, and end up getting derivative 0 since the derivative of any constant function is zero.

To be honest, I hope that none of you would make this particular error, because this example is fairly straightforward. But when you deal with more complicated relations—say, u and v are both functions of t, y is a function of u, and you have some equation that involves all four letters t, u, v, y—it can be easy to lose track of whether you are dealing with functions or numbers “underneath.” A good rule of thumb here is the following:

Rule of Thumb. *First, do all your differentiating. Then, and only then, start treating variables as numbers.*

For instance, if you are asked to find the derivative of $x^2 + 1$ at $x = 2$, you should first differentiate (obtaining $2x$) and then substitute in $x = 2$ (obtaining 4, the correct answer). Like any rule of thumb, this one has occasional exceptions. The only truly reliable way to stay out of trouble is to know what you are doing: to know, at each step of your argument, whether $x^2 + 1$ really means “the number $x^2 + 1$” or “the function that maps $x \mapsto x^2 + 1$.” However, trying to keep track of this can be quite confusing, and I think the Rule of Thumb above will probably serve you well.

Assignment 21 (due Wednesday, 23 November)

[Note: If you won't be in class on the day before Thanksgiving, then some time before class, put the homework in my mailbox (in the Eckhart basement). Also, send me an email so that I know you have done this.]

Section 2.6, Problems 11–12 and 20–21. Problems 12 and 21 will be graded carefully.

Section 2.7, Problems 3–6 and 19–20. The even-numbered problems will be graded carefully.

("Semi-bonus problem") Suppose that

$$\lim_{x \to c^-} f(x) = \ell = \lim_{x \to c^+} f(x);$$

i.e., both the one-sided limits are defined, and they are equal. Use the ε-δ definition of the limit to show that the two-sided limit is also defined and equal to ℓ, i.e., that

$$\lim_{x \to c} f(x) = \ell.$$

The technique involved should be similar to that used to give ε-δ proofs for piecewise linear functions.

This is a "semi-bonus problem" in the following sense:

- If you do not seriously attempt it, you will not receive full credit on your homework.
- If you seriously attempt it, you will receive full credit for it (although your actual homework grade will, of course, depend on the other homework problems).
- If you get it right, you will receive a bonus point on the homework.

Assignment 22 (due Monday, 28 November)

Section 2.5, Problems 19 and 20.

Section 2.6, Problems 23 and 38. Both of these will be graded carefully.

Section 2.7, Problems 8, 21, and 37. Problems 21 and 37 will be graded carefully.

Section 2.8, Problem 1.

Math 131, Lecture 24: Related Rates

Charles Staats

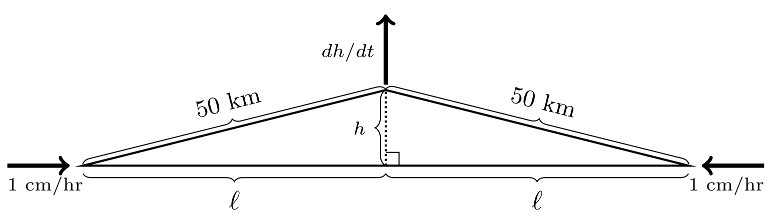
Wednesday, 23 November 2011

As far as I can tell, "related rates" are the textbook's first excuse to really start in on so-called "word problems." Up to now, the course has been mostly theoretical; the only real "applications" have been to studying the graphs of functions. However, calculus was invented for real-world problems. If you can't understand how calculus relates to the real world, then you don't really understand calculus at all.

I think the real meat of the notion of "related rates" is in the examples, so let us proceed to these examples without further ado.

Example 1. Suppose that a straight railroad consists of two completely rigid segments, each 50 kilometers long. Suppose, further, that two immensely strong men move the ends of the railroad toward each other at a constant rate of one centimeter per hour, forcing the railroad to rise up in the center. (By this, I mean that each end of the railroad is moving at a rate of one centimeter per hour.) After one hour, how fast is the center point of the railroad moving up?

Solution. In any word problem like this, the first step is almost always to draw a picture. At the same time, we probably want to assign names to all the variable quantities.



What we are interested in calculating is the rate of change of the height h with respect to time t. We need to fix units, so let's say we take time in hours and distance in kilometers. We are given that the horizontal length ℓ is shrinking at a rate of

$$1\frac{\text{cm}}{\text{hr}} = .00001\frac{\text{km}}{\text{hr}}.$$

In other words,

$$\frac{d\ell}{dt} = -.00001.$$

The Pythagorean Theorem tells us that

$$h^2 + \ell^2 = 50^2;$$

differentiating both sides with respect to t, we see that

$$\begin{aligned} 2h\frac{dh}{dt} + 2\ell\frac{d\ell}{dt} &= 0 \\ 2h\frac{dh}{dt} + 2\ell(-.00001) &= 0 \\ \frac{dh}{dt} &= \frac{.00001\ell}{h} \\ &= \frac{1}{100000h}. \end{aligned}$$

Up to now, we have only been making substitutions when we knew that something held for all time (or at least, all $t > 0$). This is because we needed to be able to differentiate; and as discussed last time, this means we are really working with functions rather than numbers. Our variables that could change over the course of time, would need to remain "dummy variables" so that we could differentiate them (or with respect to them).

However, we are done differentiating now, so we can substitute in the particular case we care about: specifically, when $t = 1$ (i.e., after one hour). In this case, we have that

$$\ell = 50 - .00001 = 49.99999.$$

Since there is an h in the formula for dh/dt, we also need to find out what h is at $t = 1$, which we do using the Pythagorean Theorem (again):

$$\begin{aligned} h^2 + \ell^2 &= 50^2 \\ h^2 &= 50^2 - \ell^2 \\ &= 50^2 - (50 - .00001)^2 \\ &= 50^2 - 50^2 + 2 \cdot 50(.00001) - .0000000001 \\ &= .0001 - .0000000001 \\ h &= \sqrt{.0001 - .0000000001} \\ &= \sqrt{10^{-4} - 10^{-10}} \\ &= \sqrt{10^{-4}(1 - 10^{-6})} \\ &= 10^{-2}\sqrt{1 - 10^{-6}}. \end{aligned}$$

Thus, plugging in this h, we find that

$$\left.\frac{dh}{dt}\right|_{t=1} = \frac{1}{10^5 h}$$
$$= \frac{1}{10^5 \cdot 10^{-2}\sqrt{1-10^{-6}}}$$
$$= \frac{10^{-5+2}}{\sqrt{1-10^{-6}}}$$
$$= \frac{.001}{\sqrt{1-10^{-6}}}.$$

Since $\sqrt{1-10^{-6}}$ is very nearly 1, this tells us that the rate of the vertex going up, dh/dt, is very close to .001 kilometers per hour, or 1 meter per hour, at time $t = 1$ hr. Thus, the midpoint is going up much faster the sides are going in (1 cm/hr). □

If you really think about it, the problem above does not so much calculate the rate of change, as explain why the problem is so incredibly unrealistic. The way to make work easier is to use leverage, or "mechanical advantage," so that your quick motion produces a slow motion in the thing you are trying to move. The fictional "very strong men" in this example are doing exactly the opposite: they are working at an enormous mechanical disadvantage.

Assignment 22 (due Monday, 28 November)

Section 2.5, Problems 19 and 20.

Section 2.6, Problems 23 and 38. Both of these will be graded carefully.

Section 2.7, Problems 8, 21, and 37. Problems 21 and 37 will be graded carefully.

Section 2.8, Problem 1.

Math 131, Lecture 25

Charles Staats

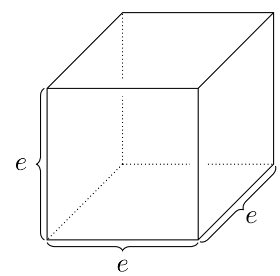
Monday, 28 November 2011

1 Related Rates examples

First, let's go over the related rates homework problem due today.

Example 1. (Section 2.8, Problem 1) Each edge of a variable cube is increasing at a rate of 3 inches per second. How fast is the volume of the cube increasing when an edge is 12 inches long?

Solution. Let e denote the edge length of the cube, and let V denote its volume.



These two quantities are related by the equation

$$V = e^3.$$

Differentiating implicitly, we see that

$$\begin{aligned}\frac{dV}{dt} &= 3e^2\frac{de}{dt}\\ &= 3e^2 \cdot 3\\ &= 9e^2.\end{aligned}$$

(We substituted in $de/dt = 3$ since this is true for all time.) At the particular instant we care about, we are given that $e = 12$, and so

$$\frac{dV}{dt} = 9e^2 = 9(12)^2 = 9 \cdot 144 = 1296.$$

$$\boxed{\text{The volume is increasing at a rate of 1296 cubic inches per second.} \qquad \square}$$

Now, another example:

Example 2. (Example 1, p. 135 in the textbook) A small balloon is released at a point 150 feet away from an observer, who is on level ground. If the balloon goes straight up at a rate of 8 feet per second, how fast is the distance from the observer to the balloon increasing when the balloon is 50 feet high?

Solution. I'm not going to type out the solution since it is explained in the text, but I will leave some space here for you to take notes on what is said in class (if you choose to do so).

□

2 Maxima and Minima

Consider a child selling lemonade on the sidewalk.[1] If she sets the price at \$0 per cup (i.e., she gives it away for free), then plenty of people will take a cup, but she won't make any money. On the other hand, if she sets the price too high—say, \$7 per cup—then no one will buy from her, and she also won't make any money. If she puts the price somewhere in the middle, then she may well sell some lemonade and make some money. But how can she figure out what price to set so that she will make the most money? Realistically, she probably can't—but only because she does not know calculus.[2]

Let m denote the amount of money she makes, let p denote the price she charges, and let n denote the number of cups she sells. It is fairly clear that

$$m = n \cdot p;$$

in words, the amount of money she makes is the number of cups she sells times the price per cup.[3] Moreover, we are assuming that the number of cups she

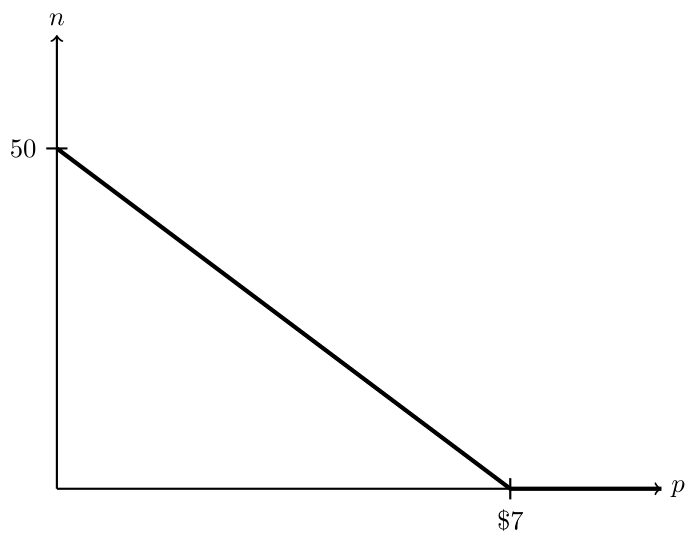
[1] Let's pretend it's summer; otherwise, she has chosen a singularly inappropriate time of year for her enterprise.

[2] Okay, I'm exaggerating here. She would also need to have done a fair amount of market research, and even then the answer would only be approximate. But since this is a course in calculus rather than economics, we're going to ignore that bit.

[3] You might object that we should also consider how much she has to pay for the lemonade, but I'm assuming her mom covers that for her.

sells is determined by the price she sets. In other words, n is a function of p. Consequently, m is also a function of p.

Ideally, we should do a fair amount of market research to figure out what function gives n; in other words, how many cups sells when she sets a given price. But since we're mathematicians rather than economists here, let's just make a sort of silly guess. Let's say that if she sets the price at \$0, then she will "sell" (give away) 50 cups (maybe 50 people pass by during the hour she sits at the stand). If she sets the price to \$7 or more, she will sell zero cups. So, let's just draw a straight line between the points $(0, 50)$ and $(7, 0)$, and call it n.



The corresponding function is

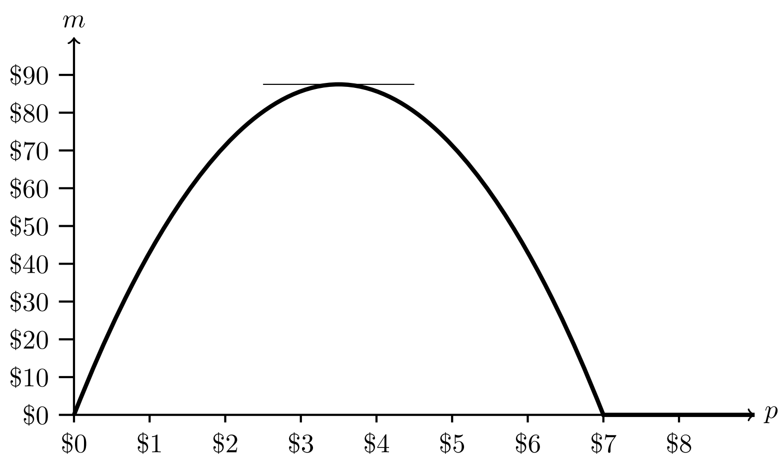
$$n = \begin{cases} 50 - \frac{50}{7}p & \text{if } 0 \leq p \leq 7, \\ 0 & \text{if } p > 7 \end{cases}$$

$$m = np = \begin{cases} \left(50 - \frac{50}{7}p\right) p & \text{if } 0 \leq p \leq 7, \\ 0 \cdot p & \text{if } p > 7 \end{cases}$$

$$= \begin{cases} 50p - \frac{50}{7}p^2 & \text{if } 0 \leq p \leq 7, \\ 0 & \text{if } p > 7. \end{cases}$$

Note that these functions are not defined for $p < 0$, since "negative price" really does not make sense in this context.

If we graph m, the amount of money made, as a function of the price p, we obtain



The maximum value is the point where the tangent line to the graph is horizontal—in other words, where $m'(p) = 0$. And we can find this using calculus:

$$m(p) = \begin{cases} 50p - \frac{50}{7}p^2 & \text{if } 0 \leq p \leq 7, \\ 0 & \text{if } p > 7. \end{cases}$$

$$m'(p) = \begin{cases} 50 - \frac{100}{7}p & \text{if } 0 < p < 7, \\ 0 & \text{if } p > 7. \end{cases}$$

Warning. *One error that a* lot *of people made on the test would amount, in this case, to writing $m'(p) = 50 - \frac{100}{7}p$ for $0 \leq p \leq 7$. (Note the $\leq$ sign rather than the $<$ sign.) When you differentiate a piecewise-defined function, a $\leq$ sign will usually (although not always) become a $<$ sign. If you look at the graph, you can see that the function is not differentiable at $p = 7$.*

If we solve for the places where $m'(p) = 0$, we find that this holds when $p = 3/2$ or $p > 7$. Looking at the graph, it is clear that m is maximized (i.e., the girl makes the most possible money) when $p = 7/2 = 3.5$; in other words, according to this model, she ought to set her price at \$3.50 per cup. The maximum value of the function is

$$m\left(\tfrac{7}{2}\right) = 50\left(\tfrac{7}{2}\right) - \tfrac{50}{7}\left(\tfrac{7}{2}\right)^2 = 87.5.$$

In other words, the most money the girl can possibly make is \$87.5.

The following, more precise mathematics allows us to handle these sorts of things more generally:

Definition. Let f be a function defined on an interval $[a, b]$ and x_0 a point in its domain. We say that x_0 is a *critical point* of f if any of the following holds:

- x_0 is an endpoint of the interval (i.e., $x_0 = a$ or $x_0 = b$); or
- $f'(x_0)$ does not exist; or

- $f'(x_0) = 0$.

The last type of critical point, where $f'(x_0) = 0$, is in some sense the most interesting sort of critical point to find (find the derivative f', then solve for $f'(x) = 0$). But the other two kinds should not be forgotten, since they are absolutely necessary to make the following theorem true.

Theorem. Let f be a *continuous* function with domain a *closed* interval $[a, b]$. Then f has a maximum value and a minimum value. Moreover, every point at which the maximum (minimum) is attained is a critical point.

In other words, if we know f is a continuous function on $[a, b]$, then the following procedure will allow us to find the minima and maxima of f on $[a, b]$:

1. Find the critical points of f (all three kinds).

2. Evaluate f at each of the critical points.

3. The largest of the resulting values is the maximum value of f on $[a, b]$. The least of the resulting values is the minimum value of f on $[a, b]$.

Assignment 23 (last assignment; due Wednesday, 30 November, 2011)

Section 2.7, Problems 21, 22, and 38.

Section 2.8, Problems 3, 6, and 17(a,b).

Section 3.1, Problems 1 and 5–6. On 5 and 6, include graphs of the function on the interval. Do NOT graph the function outside the interval.

Bonus Exercise. *Let f be a function defined on $(0,1)$; in other words, $f(x)$ is defined whenever $0 < x < 1$.*

(i) Give a formal (M-δ) definition for the statement that

$$\lim_{x \to 0^+} f(x) = \infty.$$

(ii) Assume that $\lim_{x \to 0^+} f(x) = \infty$. Use the M-δ statement above to prove that f has no maximimum value on $(0,1)$.

Math 131, Lecture 26 (final lecture)

Charles Staats

Wednesday, 30 November 2011

1 Logistics: Review session on Friday

Since reading period starts tomorrow, class on Friday will not introduce any new material. Instead, I will devote the class to answering students' questions. I expect most of the time to be spent on going over how to solve different kinds of problems, but I will also take questions about what sorts of things are and are not fair game for the exam.

Attendance is not required, but I think the class will be more helpful for everyone if a lot of people show up. I want people to do well on the final, and it is very frustrating for me when people miss something that they could have gotten right if they had only asked me to explain it.

Chaofan and Jay will not be holding tutorials tomorrow. Seth, however, will (in Pick 022); everyone is welcome to attend, whether or not you are in Seth's tutorial normally.

2 Maxima and minima—motivation applications

We will be studying how to use calculus (specifically, derivatives) to find points at which a function is maximized or minimized. This sort of thing has many practical applications. For instance, we can ask

- What price should we sell lemonade at in order to make the most profit? (Profit is a function of price; we want to select price to maximize it.)
- What shape should a rectangle be to fence in the largest possible area with a fixed amount of fence? (The area of the rectangle is a function of its length; we want to maximize it.)
- What path should a pipeline follow under a river to minimize the cost of building it? (The cost is a function of the path; we want to minimize this function.)

We won't get to solve these sorts of problems in this lecture (and thus not until next quarter), but I will show you the mathematical tools that are used to solve them.

3 Maxima and minima—the theory

We're going to spend a few minutes talking about the basic theory (theorems and such) before seeing the applications.

Definition. Let f be a function. The *maximum value* of f is a value M such that

(i) f attains the value M; i.e., there is some x_0 such that $M = f(x_0)$; and

(ii) $M \geq f(x)$ for all x in the domain of f.

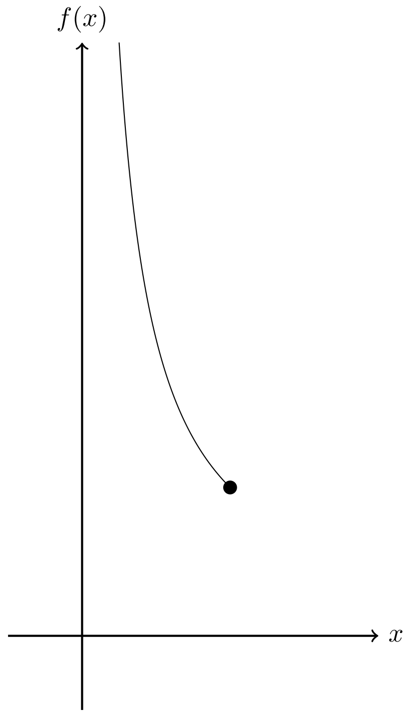
The *minimum value* of f is a value m such that

(i) f attains the value m; i.e., there is some x_0 such that $m = f(x_0)$; and

(ii) $m \leq f(x)$ for all x in the domain of f.

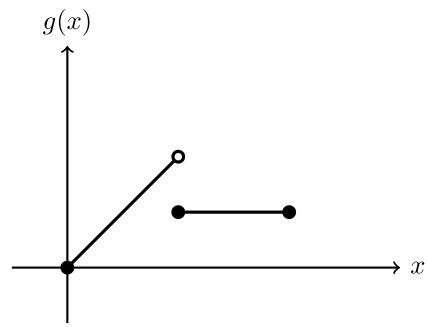
An *extreme value* of f is a value y that is either the maximum or the minimum value of f.

Warning. *Maximum and minimum values need not exist; consider the following two cases.*



$$f(x) = \frac{1}{x} \quad \textit{if } 0 < x \leq 1$$

This function has no maximum because it attains arbitrarily large values.



$$g(x) = \begin{cases} x & \textit{if } 0 \leq x < 1, \\ \frac{1}{2} & \textit{if } 1 \leq x \leq 2. \end{cases}$$

The maximum of this function "should" be 1*, but in fact the function has no maximum because it never quite reaches* 1*. There is no point* x_0 *such that* $g(x_0) = 1$*.*

In both of the cases above, the "issue" was that there were points at which the function had no finite limit. Specifically,

$$\lim_{x \to 0} f(x) = \infty, \quad \text{while} \quad \lim_{x \to 1} g(x) \text{ does not exist.}$$

This yields plausibility to the following theorem:

Theorem. Let f be a *continuous* function on a *closed* interval $[a, b]$. Then f has a minimum and a maximum.

We won't even try to prove this. For the function f above, the function was defined on $(0, 1]$, but 0 was missing from the domain—the interval was not closed. For g, the function was not continuous.

The points where minimum and maximum values might take place are called *critical points*. More precisely,

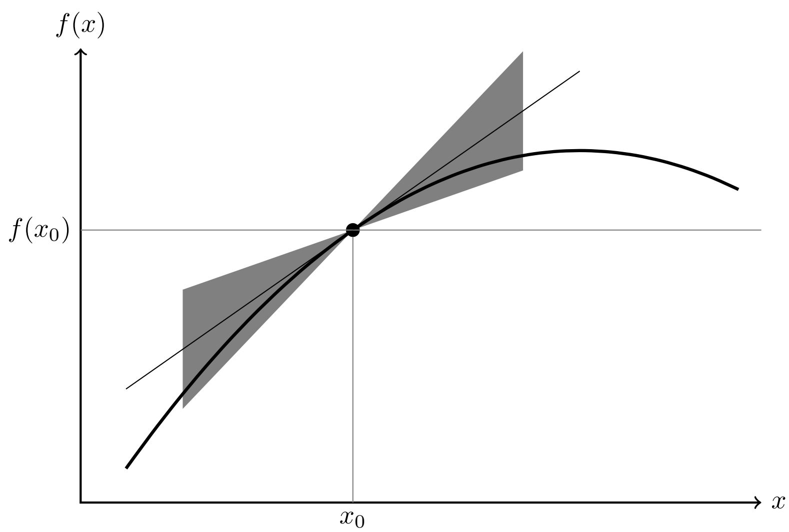
Definition. Let f be a function defined on an interval $[a, b]$ and x_0 a point in its domain. We say that x_0 is a *critical point* of f if any of the following holds:

- x_0 is an endpoint of the interval (i.e., $x_0 = a$ or $x_0 = b$); or
- $f'(x_0)$ does not exist; or
- $f'(x_0) = 0$.

The last type of critical point, where $f'(x_0) = 0$, is in some sense the most interesting sort of critical point to find (find the derivative f', then solve for $f'(x) = 0$). But the other two kinds should not be forgotten, since they are absolutely necessary to make the following theorem true.

Theorem. Let f be a *continuous* function with domain a *closed* interval $[a, b]$. Then the only points where f could possibly equal its extreme values are the critical points.

Idea of proof. We prove the contrapositive. Suppose x_0 is not a critical point. We will show that $f(x_0)$ is not an extremal value of f.



Since x_0 is not a critical point, x_0 is differentiable and $f'(x_0) \neq 0$. In other words, f has a tangent line at x_0 that is not horizontal. Thus, for x sufficiently close to x_0, $f(x)$ is contained in a narrow cone about the tangent line.

Since the tangent line is not horizontal, if we make the cone sufficiently narrow, we can ensure that the values of f immediately to the right of x_0 (if the slope is positive) or immediately to the left of x_0 (if the slope is negative) are above $f(x_0)$. Since x_0 is not a critical point, it is not an endpoint of the domain, so f does have values immediately to the left and right of x_0. Hence, $f(x_0)$ is not an maximum of f.

Similar reasoning shows that $f(x_0)$ is not a minimum value of f. □

4 Maxima and minima: example

In other words, if we know f is a continuous function on $[a, b]$, then the following procedure will allow us to find the minima and maxima of f on $[a, b]$:

1. Find the critical points of f (all three kinds).
2. Evaluate f at each of the critical points.
3. The largest of the resulting values is the maximum value of f on $[a, b]$. The least of the resulting values is the minimum value of f on $[a, b]$.

Example 1. Find the critical points, minimum, and maximum for the function f given by

$$f(x) = \tfrac{1}{3}x^3 - x$$

on the closed interval $[-2.5, 1.5]$.