MATH 131, LECTURE 1

CHARLES STAATS

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.

One of the key facts about analysis is that *there is no least positive number*. The way we prove this is called *proof by contradiction*: imagine a world in which there *were* a least positive number, and show that this world does not even make sense.

Date: 1 October 2012.

Thus, our initial assumption—that there exists a least positive number—must have been mistaken.

Before we can start on the proof, we need to clarify a couple things. First, what do we mean by "number"? If the only numbers we consider are the "whole numbers" $1, 2, 3, 4, \ldots$, then there clearly *is* a least positive number: 1. However, if we consider numbers that include also $0.1, 0.01, 0.001, 0.0001, \ldots$, the statement that "there is no least positive number" becomes more plausible.

In this course, we will deal with two kinds of numbers:

Definition. The *integers* are the numbers (including negative numbers) with no fractional part. For example, 0, 1, 2, 17, -1, -2, and -411 are all integers, but 4.1, $\sqrt{2}$, and $-\pi$ are not.

Definition. The *real numbers* are all those numbers (including positive numbers, negative numbers, and zero) that can be written with a (possibly infinite) decimal expansion. Every integer is a real number, so 0, 1, 2, 17, -1, -2, and -411 are all real numbers. But there are also real numbers like 4.1, $\sqrt{2}$, and $-\pi$ that are not integers.

The other thing that needs clarifying is what exactly is meant by "least number."

Definition. A number x is the *least number* in a collection of numbers if

- x lies in the collection AND
- for every number y in the collection, $x \leq y$.

In other words, x is the least number if x is smaller than every other number.

Now, we're ready to give the proof.

Theorem. There is no least positive real number.

Proof. Suppose, by way of contradiction, that there *were* a least positive real number. Call it x. Then $\frac{1}{2}x$ would also be a positive real number.

- Since x is the least positive real number, it is smaller than every other positive real number. In particular, $x \leq \frac{1}{2}x$.
- Since x is positive, we can multiply the inequality

$$0 < \tfrac{1}{2} < 1$$

  through by x, obtaining

$$0x < \tfrac{1}{2}x < 1x$$
$$0 < \tfrac{1}{2}x < x.$$

  Thus, $x > \frac{1}{2}x$.

We have shown that $x \leq \frac{1}{2}x$ and that $x > \frac{1}{2}x$. These cannot both be true: we have obtained a contradiction.

Therefore, our initial assumption—that there exists a least positive real number—must have been mistaken. ∎

In-class Exercise. *Show that there is no greatest negative real number.*

Proof.

□

3. Assignment 0 (due Friday, 5 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. The essays will be collected and graded (by the instructor).

Additional instructions:

(1) The essay(s) about your own past learning experience(s) are about *your* experiences. The reading assignment should help you think about the past experiences, but Prof. Brown's essay is not the subject. (On the other hand, his technique is, to some extent, the subject of the final "essay.")

(2) I encourage you to submit multiple drafts of the same essay, rather than three different essays. However, I also require that each draft should be written *without looking at the previous drafts*. [This is something of an experiment; let me know how it goes.]

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

MATH 131, LECTURE 2

CHARLES STAATS

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 nonnegative 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:

Date: 3 October 2012.

> "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. I've distributed a handout of rules that you should use for reference. I am also going to draw pictures showing how these operations work on the number line, which may help you understand why the inequality sign is reversed in some cases, but not others.

- Rule 1:

- Rules 2 and 3:

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.

At this point, the purely algebraic route would be to consider four cases separately, and see which numbers are encompassed by each case:

However, it turns out that looking at a more geometric picture, involving a number line, can help to elucidate the process:

The number line suggests that we only need to consider three cases:

$$x < -2$$
$$-2 < x < 3$$
$$3 < x$$

This is a "shortcut" to seeing certain redundancies: for instance, if $x < -2$, then we don't have to specify whether x is less than 3—we already know.

0pt50pt

0pt50pt

Assignment 0 (due Friday, 5 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. The essays will be collected and graded (by the instructor).

Additional instructions:

(1) The essay(s) about your own past learning experience(s) are about *your* experiences. The reading assignment should help you think about the past experiences, but Prof. Brown's essay is not the subject. (On the other hand, his technique is, to some extent, the subject of the final "essay.")

(2) I encourage you to submit multiple drafts of the same essay, rather than three different essays. However, I also require that each draft should be written *without looking at the previous drafts*. [This is something of an experiment; let me know how it goes.]

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

Assignment 1 (due Monday, 9 October)

Read "A Bit of Logic" and "Quantifiers" on pp. 4–6.
Problem Set 0.1, numbers 27, 28, 45, 46, 63, and 64. Problems 27, 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 a* positive *real number x are equivalent:*

(i) $x < \sqrt{2}$.
(ii) $x^2 < 2$.
(iii) There exists y such that ($x < y$ AND $y^2 < 2$). *(Hint: use the Theorem on page* **??***.)*

MATH 131, LECTURE 3

CHARLES STAATS

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. Sets and interval notation

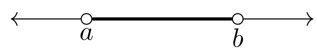
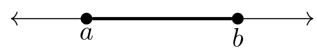
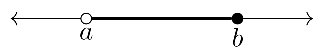
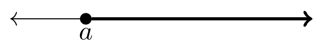
When we have a condition on numbers, we can consider the *set* of all numbers that satisfy this condition. Thus, for instance,

$$\{x \mid x \leq 3\}$$

means "the set of all numbers x fulfilling the condition that $x \leq 3$," or more simply, "the set of all numbers ≤ 3."

There are certain important sets that are commonly denoted using what is called interval notation.

Date: 5 October 2012.

Notation	Meaning	Picture
(a,b)	$\{x \mid a < x < b\}$	
$[a,b]$	$\{x \mid a \leq x \leq b\}$	
$(a,b]$		
$[a,\infty)$		

To go from a set back to the condition, we use the symbol $\in$, read "is an element of" or more simply "in."

Example. Write the condition

$$x \in [-1, 4)$$

without using any set-theoretic notation.

Solution. $-1 \leq x < 4$ $\square$

Two conditions are *equivalent* if and only if the corresponding sets are *equal*.

Warning. *There is an important semantic distinction between a* condition *and the corresponding* set*. A condition on x is either true or false, depending on the value of x. A set cannot be true or false—it simply is.*

In the other direction, two sets are equal *if they contain exactly the same numbers. Two conditions can be equivalent, but it does not really make sense to ask whether they are equal.*

In short: A set is a kind of mathematical object. A condition is a way of talking about mathematical objects. These two concepts, like the words eat *and* food*, are closely related but cannot be used interchangeably.*

3. 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.*

4. 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. Hopefully, my pictures last lecture explaining certain inequality rules give examples of how this can work.

The basic approach to graphing functions is quite simple:

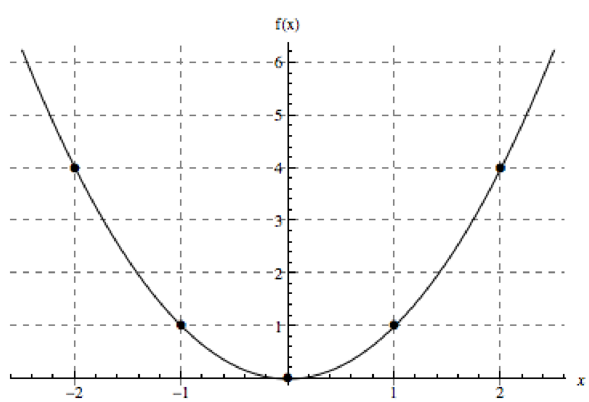
(1) Choose some values of x.
(2) Calculate and plot the points $(x, f(x))$.
(3) "Connect the dots."

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! □

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,

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

□

Assignment 1 (due Monday, 9 October)

Read “A Bit of Logic” and “Quantifiers” on pp. 4–6.
Problem Set 0.1, numbers 27, 28, 45, 46, 63, and 64. Problems 27, 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 a* positive *real number x are equivalent:*

(i) $x < \sqrt{2}$.
(ii) $x^2 < 2$.
(iii) There exists y such that ($x < y$ AND $y^2 < 2$). *(Hint: use the Theorem on the first page of Lecture 2.)*

Assignment 2 (due Wednesday, 11 October)

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

1. Clarification on implication

Some people last lecture found the implication

$$x > 3 \implies x \geq 3$$

confusing. I suspect that their confusion may have stemmed from thinking that "if $x \geq 3$, then x might be 3." Stringing these two implications together would give

$$x > 3 \implies x \geq 3 \implies x \text{ might be } 3.$$

Thus, we seem to have shown that if $x > 3$, then x might be 3, which is absurd.

The error here is the usage of "might be." A mathematical statement like $x = 3$ is either true or false, even when we don't know which. We might say that a statement "might be true" if we are, e.g., describing the narrative structure of a proof. But using the sentence "x might be 3" as an actual mathematical statement in a string of implications makes no sense: either x is equal to 3, or it's not.

For an example in real life, consider:

- If it is January, then it is not February.
- If it is not February, then it might be March.

Each of these seems plausible. But if we string them together, we get

- If it is January, then it might be March,

which is an absurd statement. The error is in including the verb "might be" in a logical statement: such statements are either true or false.

It may be helpful to remember the slogan

"*Is* or *is not*; there is no 'might be.' "

2. 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 1. (Example 13 in the book.) Describe the solution set for the inequality $x^2 - 2x - 4 < 0$. Do not use the quadratic formula.

Date: 8 October 2012.

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,

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

Hence, the solution set is $(1 - \sqrt{5}, 1 + \sqrt{5})$. □

3. 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.*

Assignment 2 (due Wednesday, 10 October)

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.

Assignment 3 (due Friday, 12 October)

Section 0.2, problems 49 and 51. Both of these 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

1. 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 $g \circ f$.

Date: 10 October 2012.

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:

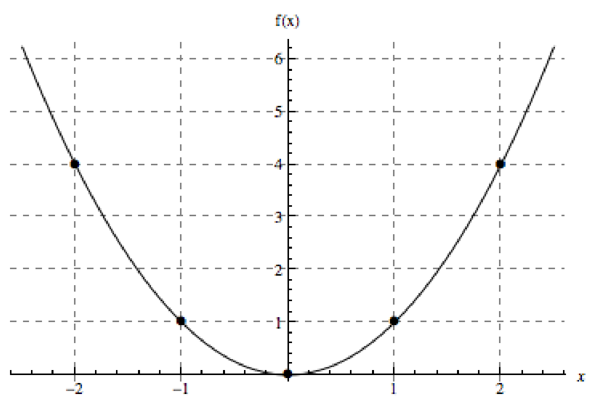
(1) Choose some values of x.
(2) Calculate and plot the points $(x, f(x))$.
(3) "Connect the dots."

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.

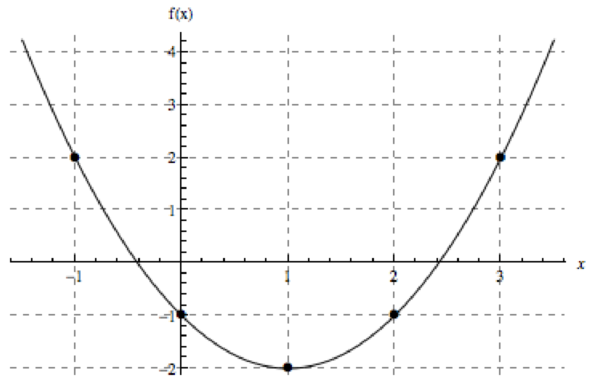
Things that can go right

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.



∎

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

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

where

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

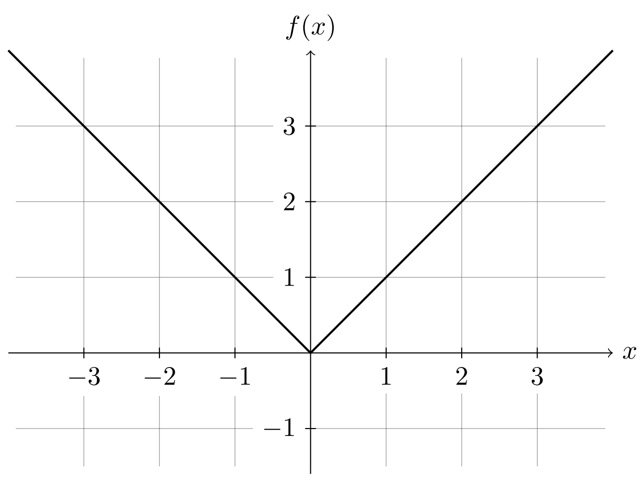
Graph the functions $g(x) = f(2x)$ and $h(x) = x + 1$. Use the resulting graph to study the set of values of x satisfying the inequality $g(x) \leq h(x)$.

3. Working with absolute values

Recall the absolute value function,

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

The graph of this function is



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.

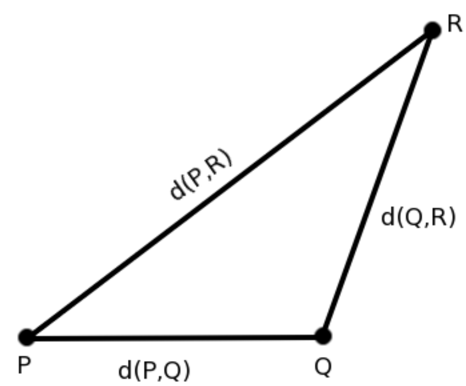
The rules for multiplication and division are easy:

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

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

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

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

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

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

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

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

$$c - a = \alpha + \beta$$
$$b - a = \alpha$$
$$c - b = \beta.$$

Thus, the triangle inequality gives us

$$|\alpha + \beta| \le |\alpha| + |\beta|,$$

which is the addition rule.

Thinking about absolute values as "distance" can also be helpful in figuring out how to simplify absolute value inequalities.

4. Quantifiers

Recall the two inequalities

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

I'm about to show how the Addition Rule implies the Subtraction Rule. I'll also use this as an excuse to discuss quantifiers.

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

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

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

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

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

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

□

Let's review the logic here. The Subtraction Rule has the appearance of a condition on a and b: that $|a+b| \geq |a| - |b|$. 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." For instance,

> 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

$$\exists y > 0 \text{ such that } y < x,$$

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

(1) Here's a specific number $y > 0$ that I've dreamed up. For instance, $y = \frac{1}{2}x$.
(2) Here's why this specific y satisfies $y < x$.

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.

Assignment 3 (due Friday, 12 October)

Section 0.2, problems 49 and 51. Both of these 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.

Assignment 4 (due Monday, 15 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.
(b) By completing the square.

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. $x^2 + 2x - 3 > 0$

Problems 2 and 3 will be graded carefully.

Lecture 6 is omitted, primarily because the lecture notes ended up including no material that actually needs to be studied in this course.

MATH 131, LECTURE 7

INSTRUCTOR: CHARLES STAATS

1. Test next week (week 4)

This is a head's up that there will be a test during Week 4, although I do not yet have many details about what will be covered and on which day it will take place.

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:

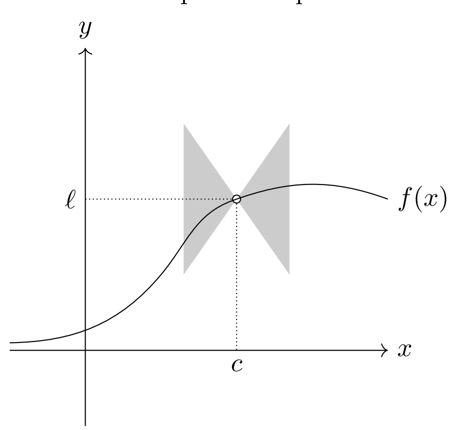
Definition. (Terribly Imprecise Version) We say that

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

if the following holds:

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

First, let's see how the "bow tie" picture captures this:



If there exists a bow tie about the point (c, ℓ) containing the graph of f, then $f(x)$ is forced to become closer to ℓ as x becomes closer to c.

The "real" definition, you may recall, is the following:

Definition. We say that

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

if the following holds:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \text{ s.t. } (|x - c| < \delta \text{ and } x \neq c) \implies |f(x) - \ell| < \varepsilon.$$

Date: 15 October 2012.

We can imagine this as a "game" as follows:

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

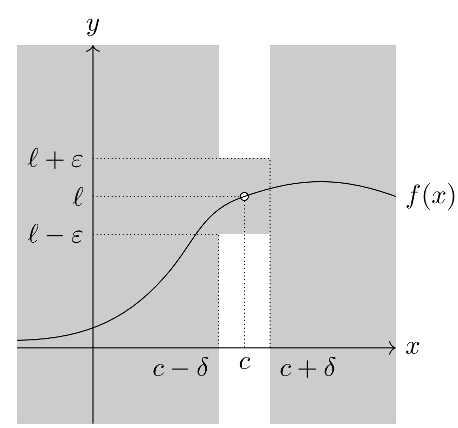
This relates to the intuitive notion, "If x is close to c then $f(x)$ is close to ℓ," because both of the "moves" amount to choosing a definition for *close to*.

(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.

The technical definition thus amounts to the following:

> No matter how we define "$f(x)$ is close to ℓ," there is a definition of "x is close to c" such that when x is close to c, then $f(x)$ is close to ℓ.

For a specific δ and ε, the picture looks like this:

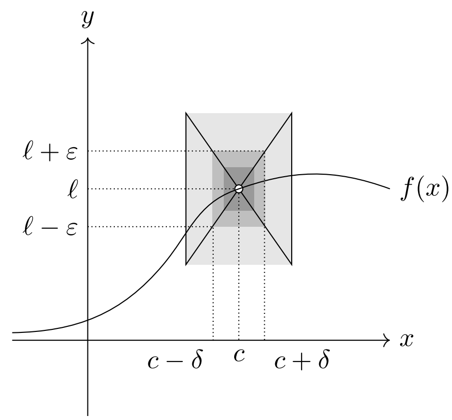


The statement

$$(|x-c| < \delta \text{ and } x \neq c) \implies |f(x) - \ell| < \varepsilon$$

holds true as long as the function remains within the shaded area. We want to ensure that this works, i.e., that the "judge rules in our favor."

Here's the relation between this and the bow tie thing: the bow tie determines acceptable "railroad tracks" for the corners of the ε-δ box:



3. Assignment 5 (due Wednesday, 17 October)

Complete the attached worksheet on graphing piecewise-defined functions. I suppose these will all be graded carefully, if only because it is easy to look at a graph and see whether it is correct.

Additionally:

- On Problem 2 of the worksheet, draw a Lipschitz bow tie about the point $(2, 3)$ that contains the function's graph where it should. Conclude that the Lipschitz limit of $f(x)$ as $x \to 2$ is 3.
- On Problem 1 of the worksheet, make a good effort to draw a Lipschitz bow tie about the point $(0, -1)$. Indicate (by circling or highlighting) the part of the graph of f that should be contained in this bow tie, but is not. (Since it is not true that $\lim_{x\to 0} f(x) = -1$, no bow tie you draw will actually work; but you should at least be able to make it work on the right-hand side.)

Again, these will all be graded carefully.

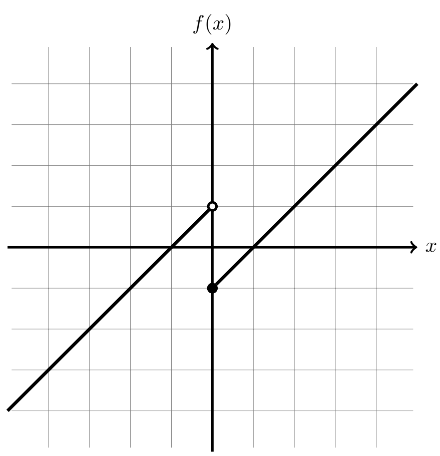
Worksheet: Graphing piecewise-defined functions

Math 131, Section 42

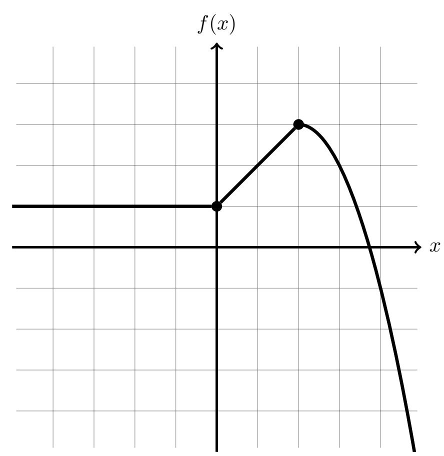
Due Wednesday, 17 October, 2011

Please graph the following functions. The first two are done for you.

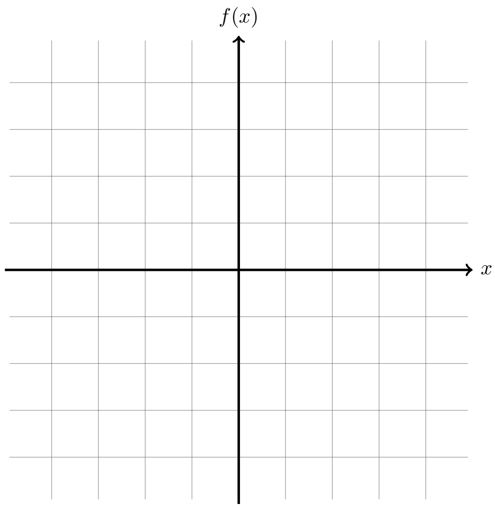
1. $f(x) = \begin{cases} x+1, & \text{if } x < 0, \\ x-1, & \text{if } x \geq 0. \end{cases}$



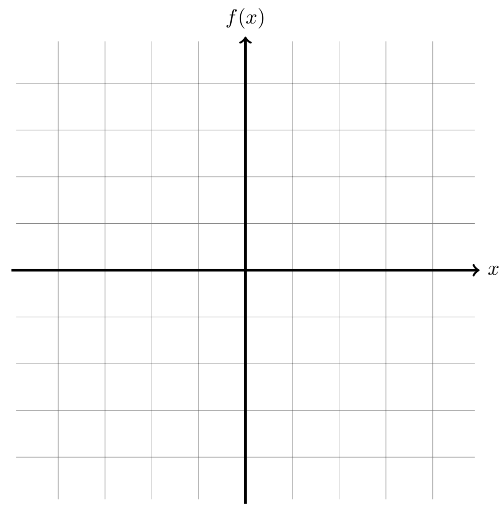
2. $f(x) = \begin{cases} 1 & \text{if } x \leq 0, \\ x+1 & \text{if } 0 < x < 2, \\ -(x-2)^2+3 & \text{if } 2 \leq x. \end{cases}$



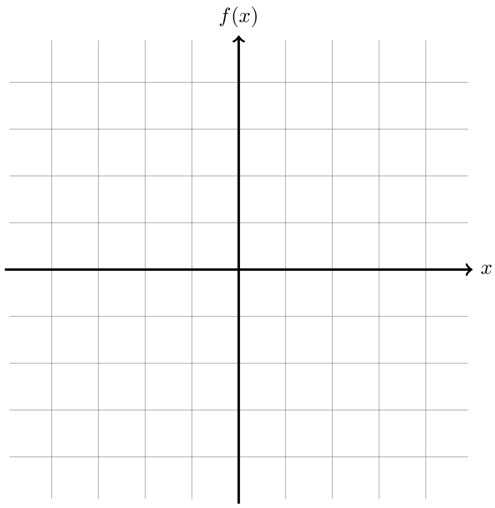
3. $f(x) = \begin{cases} 3 & \text{if } x \leq -1.5, \\ 2 & \text{if } -1.5 < x < -1.25, \\ 1 & \text{if } -1.25 \leq x < 1.25, \\ 2 & \text{if } 1.25 \leq x < 1.5, \\ 3 & \text{if } 1.5 \leq x. \end{cases}$



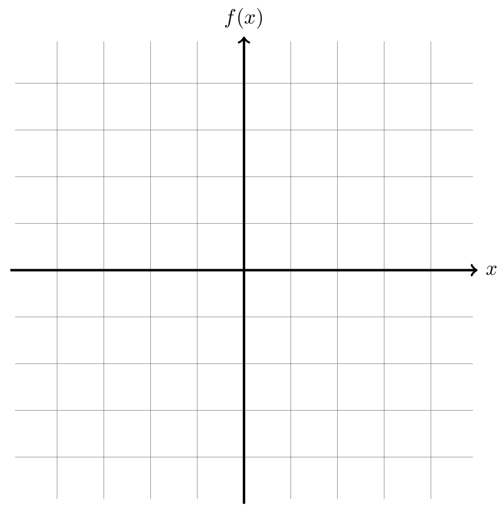
4. $f(x) = \begin{cases} 5x+7, & \text{if } x \leq -2, \\ 2x+1, & \text{if } -2 < x \leq 0, \\ 1-2x, & \text{if } 0 < x \leq 2, \\ 2, & \text{if } 2 < x. \end{cases}$



5. $f(x) = \begin{cases} -x^2 - 4 & \text{if } x \leq 1, \\ x + 2 & \text{if } x > 1. \end{cases}$

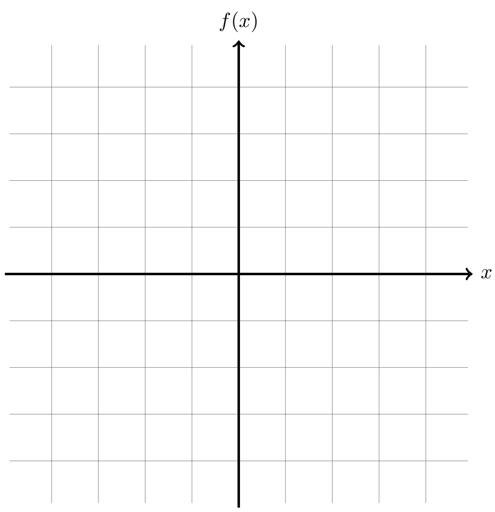


6. $f(x) = \begin{cases} (x+1)^2 & \text{if } x < 0, \\ x^2 & \text{if } 0 \leq x \leq 1, \\ 2 - 2x & \text{if } 1 < x. \end{cases}$



7. Now, you supply an interesting piecewise-defined function and graph it.

$$f(x) = \begin{cases} \qquad\qquad & \text{if} \\ \qquad\qquad & \text{if} \\ \qquad\qquad & \text{if} \end{cases}$$



MATH 131, LECTURE 8

INSTRUCTOR: CHARLES STAATS

1. Announcements

- **Test Friday, 26 October**, that is, a week from Friday.
- No office hours this Friday afternoon, since I will be headed to a conference at the University of Utah this weekend.

2. Notes from the quiz

First, some of you may be surprised, when you get back your quiz, to see how many points I may deduct even when you get the "correct 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.

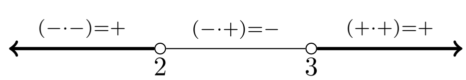
Moreover, I'm also expecting you to show me that you know how to do everything involved. You might think in terms of a "rough draft" versus a "final draft." In the "rough draft," your goal is to figure out the answer using all the techniques at your disposal. In the "final draft," you are trying to write a mathematically precise explanation for how to get to the right answer, in such a way that you are absolutely confident it is the right answer. I don't expect "final drafts" written on a quiz or a test to be polished, but I do expect to see proof that the student understands why what they are doing works.

One mistake I made, in giving instructions, has been that I have told you it was okay to use "geometric shorthand," without explaining precisely what this is shorthand for. Here's an example:

Example. Describe the solution set for the inequality

$$(x-3)(x-2) > 0.$$

Solution. **geometric shorthand:**



What this is "shorthand" for:

Divide the number line into three regions: $(-\infty, 2) \cup [2,3] \cup (3,\infty)$. Every real number x lies in exactly one of these regions.

Date: 17 October 2012.

1. If $x \in (-\infty, 2)$, then

$$\text{(a)} \qquad \begin{aligned} x - 3 &< 0 \qquad (-) \\ x - 2 &< 0 \qquad (-) \end{aligned}$$

Consequently, the product is greater than zero ($- \cdot - = +$), and so x is a solution.

2. If $x \in [2, 3]$, then

$$\text{(b)} \qquad \begin{aligned} x - 3 &\leq 0 \qquad (-) \\ x - 2 &\geq 0 \qquad (+) \end{aligned}$$

Consequently, the product is ≤ 0 ($- \cdot + = -$) and so x is *not* a solution.

3. If $x \in (3, \infty)$, then

$$\text{(c)} \qquad \begin{aligned} x - 3 &> 0 \qquad (+) \\ x - 2 &> 0 \qquad (+) \end{aligned}$$

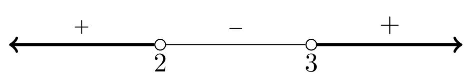
Consequently, the product is greater than zero ($+ \cdot + = +$), and so x *is* a solution.

Therefore, x satisfies the inequality if and only if $x \in (-\infty, 2) \cup (3, \infty)$.

What the "shorthand" omits: Proofs for the inequalities (a), (b), and (c). If I ask for a complete, rigorous proof, you should supply these proofs as well. □

If I let you to use the geometric shorthand, then I am, in essence, telling you that you do not have to prove statements like (a), (b), and (c).

Inadequate shorthand: If you write, as many people did,



then you are giving shorthand for the following:

Divide the number line into three regions: $(-\infty, 2) \cup [2, 3] \cup (3, \infty)$. Every real number x lies in exactly one of these regions.

1. If $x \in (-\infty, 2)$, then

$$\text{(d)} \qquad (x - 3)(x - 2) > 0,$$

and so x is a solution.

2. If $x \in [2, 3]$, then

$$\text{(e)} \qquad (x - 3)(x - 2) \leq 0,$$

and so x is *not* a solution.

3. If $x \in (3, \infty)$, then

$$\text{(f)} \qquad (x - 3)(x - 2) > 0,$$

and so x *is* a solution.

Therefore, x satisfies the inequality if and only if $x \in (-\infty, 2) \cup (3, \infty)$.

Note that you are not giving any indication of how to prove inequalities (d), (e), and (f). At this stage, I do not trust that you know how to prove these unless you give some kind of explanation, if only something like $+ \cdot - = -$.

Another point that people found confusing was the following:

$$x^2 \le 5 \iff |x| \le \sqrt{5} \iff x \in [-\sqrt{5}, \sqrt{5}]$$
$$x^2 \ge 5 \iff |x| \ge \sqrt{5} \iff x \in (-\infty, -\sqrt{5}] \cup [\sqrt{5}, \infty).$$

Some people wrote something like $x \le \pm\sqrt{5}$, which makes no sense when dealing with inequalities. (This is a device used for equations only.) Others wrote simply $x \le \sqrt{5}$, which may be in part my fault: in the inequality sheet I handed out, I included a rule that

$$x^2 \le y^2 \iff x \le y, \quad \text{assuming } x \text{ and } y \text{ are nonnegative.}$$

I made this assumption because I did not want to have to deal with absolute values on the first day of class. However, I can understand why you might now find this confusing. I'll supply a revised version of the inequality worksheet before too much longer.

3. Using Lipschitz bow ties to obtain rules for δ in terms of ε

Suppose we have drawn a Lipschitz bow tie of width $2D$ and height $2E$, both positive. Then the following rule should work for choosing δ once ε is given:

- If $\varepsilon > E$, set $\delta = D$.
- If $\varepsilon \le E$, set $\delta = \left(\dfrac{D}{E}\right)\varepsilon$.

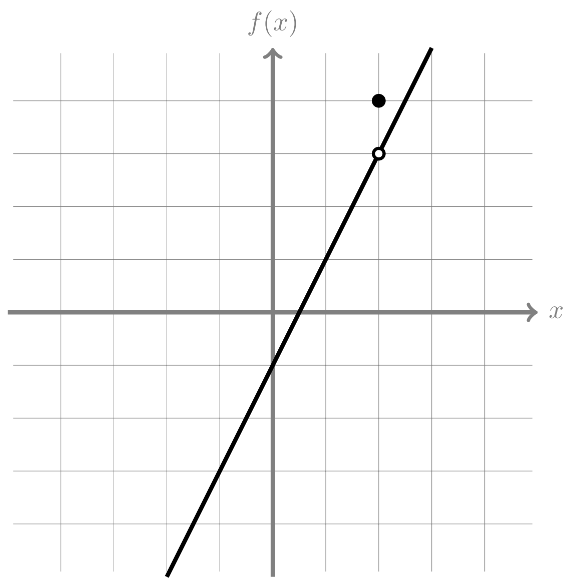
4. Examples: Using the ε-δ definition

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.

A bow tie with $D = 1, E = 3$ will work, which suggests we should use the following rule:

$$\delta = \begin{cases} 1 & \text{if } \varepsilon > 3, \\ \frac{1}{3}\varepsilon & \text{if } \varepsilon \leq 3. \end{cases}$$

Solution. Let $\varepsilon > 0$ be given.

Case 1: $\varepsilon > 3$. Set $\delta = 1$. If $|x-2| < \delta$ and $x \neq 2$, then

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

as desired.

Case 2: $0 < \varepsilon \leq 3$. Set $\delta = \frac{1}{3}\varepsilon$. Assume $|x-2| < \delta$ and $x \neq 2$.
Since $x \neq 2$, we know $f(x) = 2x - 1$. Hence,

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

as desired.

Therefore, $\lim x \to 2 f(x) = 3$. $\square$

Assignment 6 (due Friday, 19 October)

Skim Section 1.1 of the textbook.

Section 1.1, Problem 30. This is, unusually, of the "answers only" variety. Parts (c),(d), and (f) will be graded carefully.

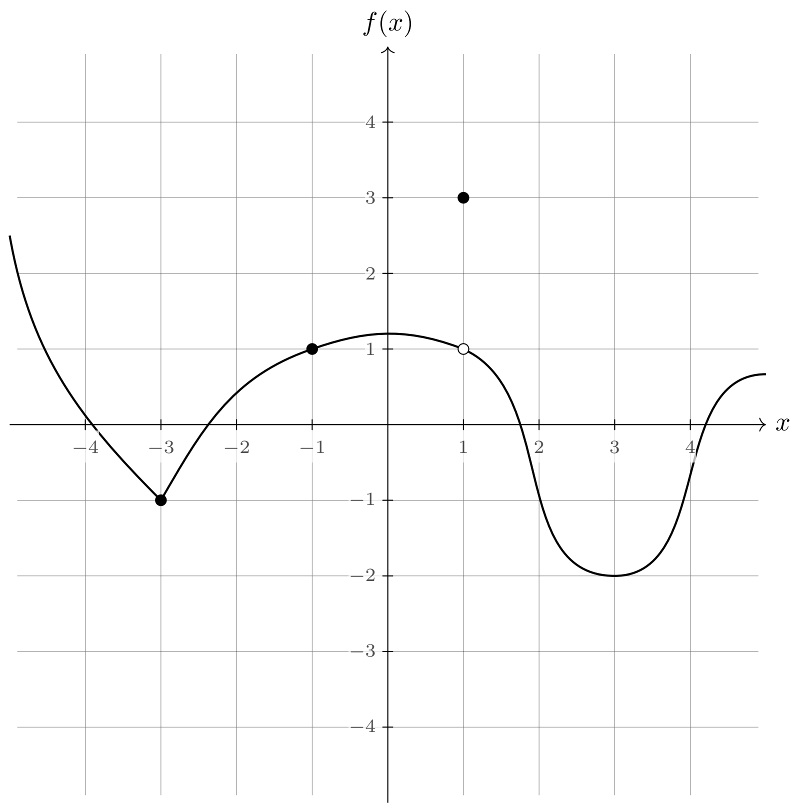
Section 1.2, Problems 11 and 12. Hint: $\delta = \frac{1}{3}\varepsilon$ will work for both of these problems, if you want to skip the "preliminary analysis." Problem 11 will be graded carefully.

In the graph below,

(a) determine, by "eyeing" the graph, the values for

$$\lim_{x \to -3} f(x), \quad \lim_{x \to -1} f(x), \quad \text{and } \lim_{x \to 1} f(x).$$

(b) At each of these three limits, draw a Lipschitz bow tie and use it to give a valid "rule" for ε depending on δ. Note: You will need to draw the bow tie carefully to accomplish this, and you may want to put its corners on grid points so that you know exactly how long and how wide it is.



MATH 131, LECTURE 9

INSTRUCTOR: CHARLES STAATS

1. Announcements

- **Test Friday, 26 October**, that is, a week from today. The test will cover the material from Lectures 1–9 (i.e., through today's lecture) and Assignments 1–8 (i.e., all mathematical assignments up to and including the one due Wednesday).
- No office hours this Friday afternoon, since I will be headed to a conference at the University of Utah this weekend.

2. 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. $$\lim_{x \to c} k = k$$
2. $$\lim_{x \to c} x = c$$
3. $$\lim_{x \to c} kf(x) = k \lim_{x \to c} f(x)$$
4. $$\lim_{x \to c} [f(x) + g(x)] = \left[\lim_{x \to c} f(x)\right] + \left[\lim_{x \to c} g(x)\right]$$
5. $$\lim_{x \to c} [f(x) - g(x)] = \left[\lim_{x \to c} f(x)\right] - \left[\lim_{x \to c} g(x)\right]$$
6. $$\lim_{x \to c} [f(x) \cdot g(x)] = \left[\lim_{x \to c} f(x)\right] \cdot \left[\lim_{x \to c} g(x)\right]$$
7. $$\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{\lim_{x \to c} f(x)}{\lim_{x \to c} g(x)}$$

Date: 19 October 2012.

8. $$\lim_{x\to c}[f(x)]^n = \left[\lim_{x\to c} f(x)\right]^n$$

"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.*

Assignment 7 (due Monday, 22 October)

- Give ε-δ proofs of the following facts:

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

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

  They will both be graded carefully.
- Section 1.3, Problems 1, 2, 14, and 15. Follow the instructions. The even-numbered problems will be graded carefully.
- In the attached handout, each of the three graphs has one good Lipschitz bow tie and one bad one. Identify which is which. For the good bow ties, identify the limit, and state the resulting rule for δ in terms of ε that could be used in an ε-δ proof. All of these will be graded carefully.

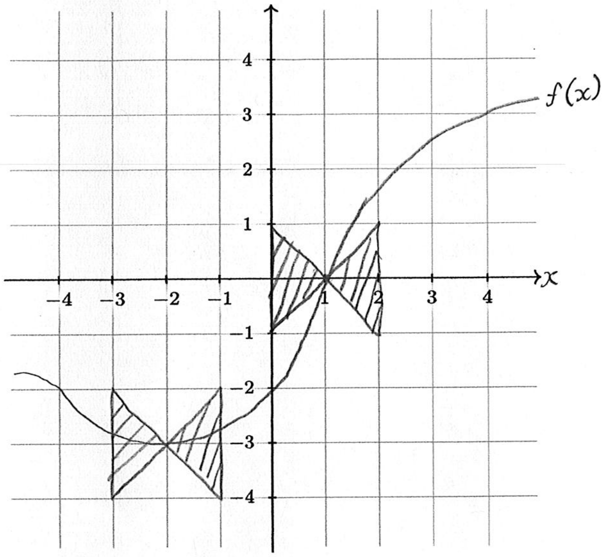
Assignment 8 (due Wednesday, 24 October)

- Give ε-δ proofs of the following facts:

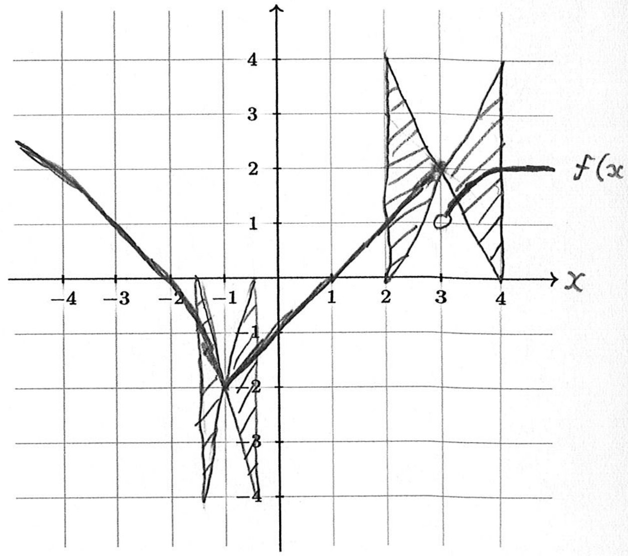
$$\lim_{x \to -\frac{1}{2}} 4x + 1 = -1 \tag{3}$$

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

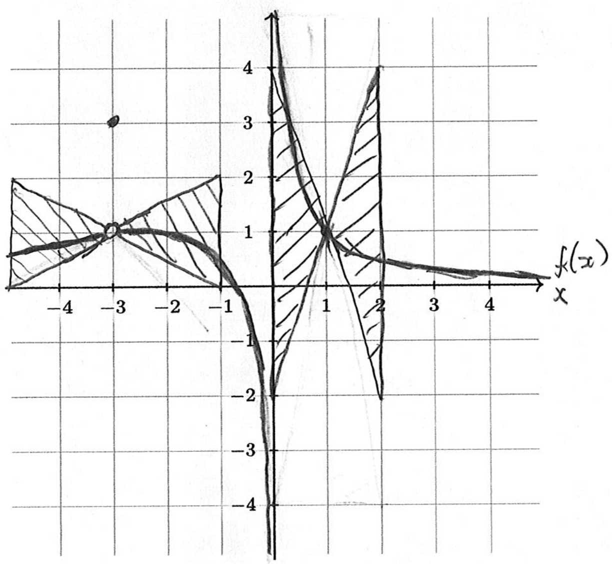
  They will both be graded carefully.
- Section 1.3, Problems 5, 6, 19, and 21. Follow the instructions. The even-numbered problems will be graded carefully.
- Section 1.6, Problem 13.



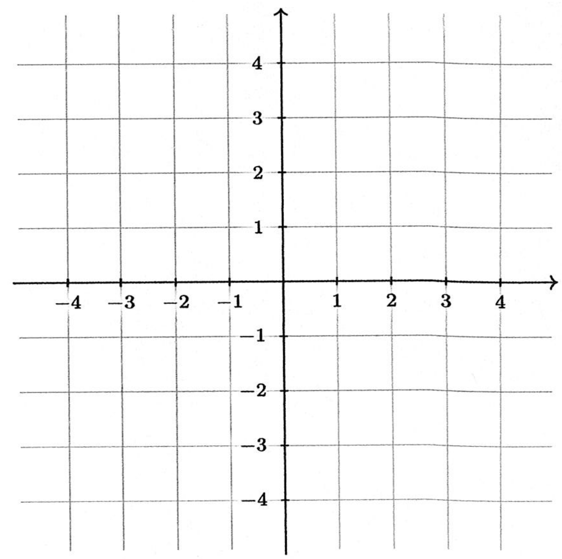
f(x)
x

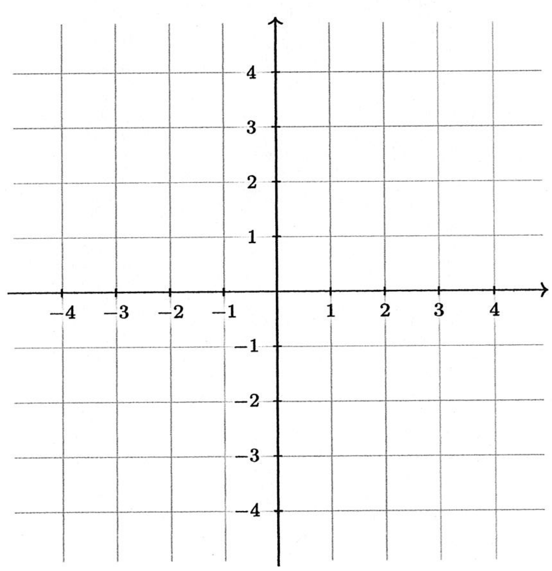


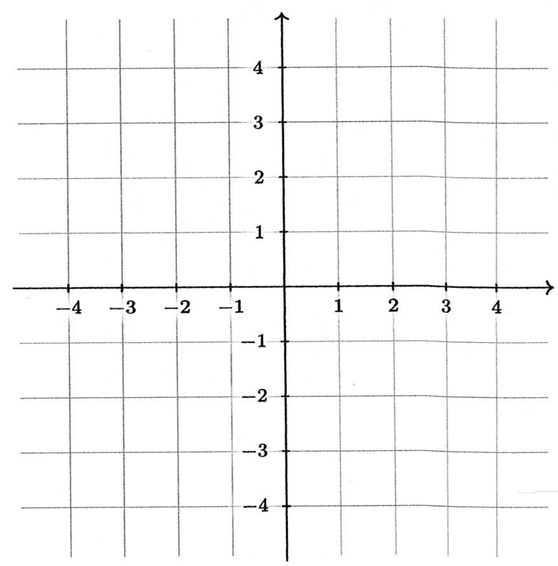
f(x)
x



f(x)
x







MATH 131, LECTURE 10

INSTRUCTOR: CHARLES STAATS

1. Announcements

Test Friday, 26 October, that is, a week from today. The test will cover the material from Lectures 1–9 (i.e., through today's lecture) and Assignments 1–8 (i.e., all mathematical assignments up to and including the one due Wednesday).

Promise: I will have a careful writeup of exactly what a "Lipschitz bow tie" is, in mathematical terms, written and posted on Chalk by 10pm tonight, October 22, 2012. Studying from it may not be the best way to study, but I think it should at least be there.

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

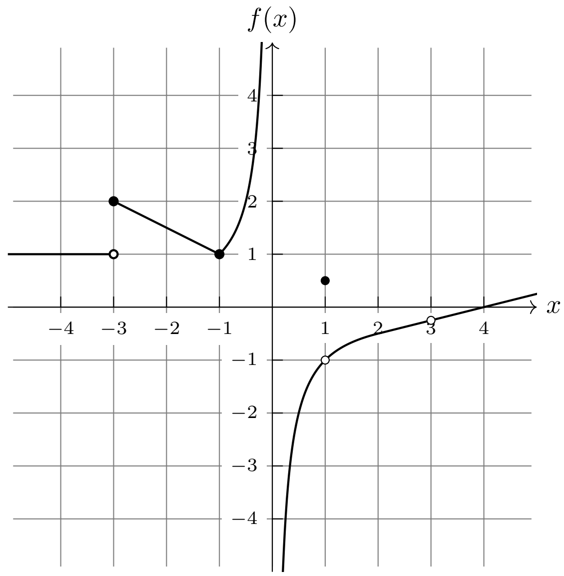
(i) f is defined at x_0, AND
(ii) $\lim_{x \to x_0} f(x)$ exists, AND
(iii) $\lim_{x \to x_0} f(x) = f(x_0)$.

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.*

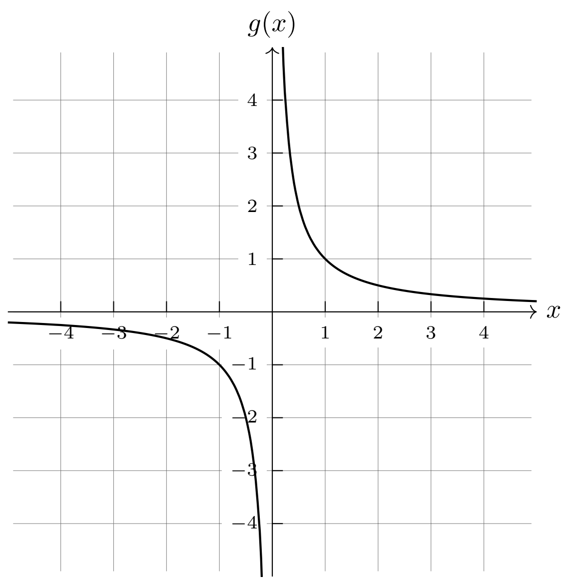
Example 1. Consider the function f whose graph looks like this:

Date: 22 October 2012.



(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 2. 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.]

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 3. 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,

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

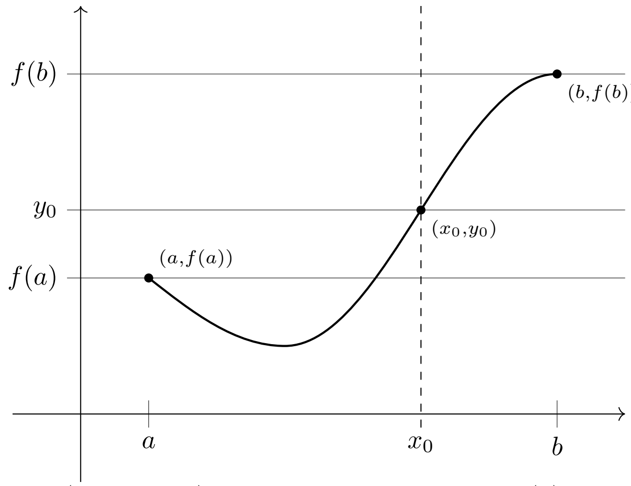
In other words, f is continuous. □

3. The intermediate value theorem

One intuitive notion of continuity 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 for defining 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:

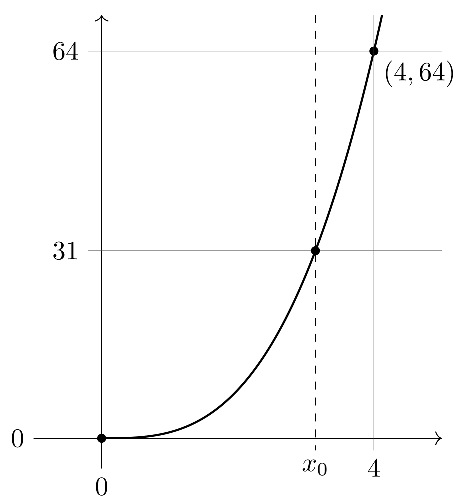


Intuitively, for the (continuous) f to go from the line $y = f(a)$ to the line $y = f(b)$, it has to pass through the line $y = y_0$ somewhere. That "somewhere" is our x_0.

We won't try to prove this right now.

Example 4. Show that $\sqrt[3]{31}$ exists. In other words, show that there is a positive real number x_0 such that $x_0^3 = 31$.

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 8 (due Wednesday, 24 October)

- Give ε-δ proofs of the following facts:

(1) $$\lim_{x \to -\frac{1}{2}} 4x + 1 = -1$$

(2) $$\lim_{x \to 5} \tfrac{1}{2}x - 2 = \tfrac{1}{2}$$

  They will both be graded carefully.

- Section 1.3, Problems 5, 6, 19, and 21. Follow the instructions. The even-numbered problems will be graded carefully.
- Section 1.6, Problem 13.

Test Friday, 26 October

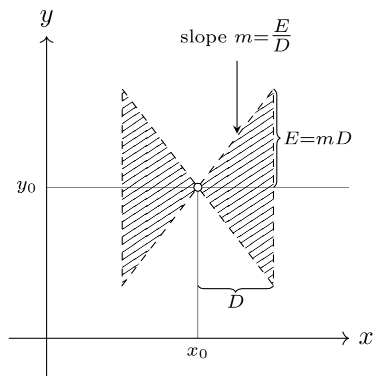
LIPSCHITZ BOW TIES

INSTRUCTOR: CHARLES STAATS III

A bow tie shaped region about the point (x_0, y_0) is specified by its half-width D and half-height E as

$$\left\{(x,y) : |x - x_0| < D \text{ AND } |y - y_0| < \tfrac{E}{D}|x - x_0|\right\}.$$

One may also specify D together with the slope m; then E is given by $E = mD$.

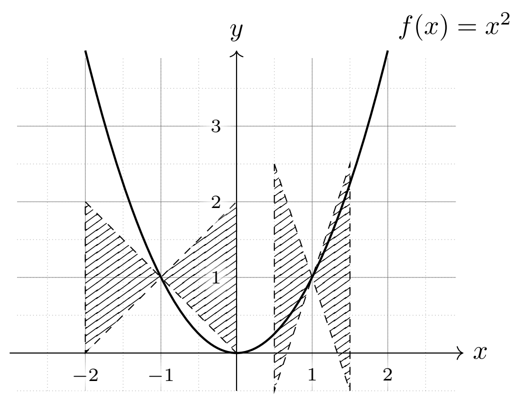


Definition. For a given function f and a point (a, ℓ), the bow tie shaped region specified by D and m is a *Lipschitz bow tie* for f if

$$|x - a| < D \text{ AND } x \neq a \implies |f(x) - \ell| < m|x - a|.$$

In other words, the bow tie is Lipschitz if the graph of $f(x)$ lies within the bow tie whenever x is within D of a (excluding $x = a$).

Example 1. In the graph below, the bow tie shaped region on the right ($D = \frac{1}{2}$, $E = \frac{3}{2}$, $m = 3$) is a Lipschitz bow tie for f; the one on the left ($D = 1$, $E = 1$, $m = 1$) is not.



Date: November 5, 2012.

Definition. Let f be a function whose domain includes $(a-d,a)\cup(a,a+d)$. The statement "ℓ is the Lipschitz limit of $f(x)$ as $x\to a$," written in symbols as

$$\operatorname*{Lip\,lim}_{x\to a} f(x) = \ell,$$

is defined to mean the following:

$$\exists D>0, E>0 \text{ s.t. if } x\in(a-D,a)\cup(a,a+D), \text{ then } |f(x)-\ell| < \left(\frac{E}{D}\right)|x-a|.$$

Let's rewrite this with some explanation:

$$\exists \underbrace{D>0}_{\substack{\text{half-width}\\ \text{of bow tie}}}, \underbrace{E>0}_{\substack{\text{half-height}\\ \text{of bow tie}}} \text{ s.t. if } \underbrace{x\in(a-D,a)\cup(a,a+D)}_{x\text{ is within } D \text{ of } a}, \text{ then } \underbrace{|f(x)-\ell| < \left(\tfrac{E}{D}\right)|x-a|}_{f(x)\text{ lies within the bow tie}}$$

In other words, $\operatorname{Lip\,lim}_{x\to a} f(x)=\ell$ if and only if there exists a Lipschitz bow tie for f about (a,ℓ). Thus, for instance, in Example 1, the Lipschitz bow tie on the right shows that

$$\operatorname*{Lip\,lim}_{x\to 1} x^2 = 1.$$

The bow tie on the left does *not* show that

$$\operatorname*{Lip\,lim}_{x\to -1} x^2 = 1,$$

because it is not Lipschitz for $f(x)=x^2$. It is nevertheless true that $\operatorname{Lip\,lim}_{x\to-1} x^2 = 1$; even though the bow tie chosen does not work, in this case some other bow tie would be Lipschitz.

Exercise 2. Draw a "better" bow tie about $(-1,1)$ on the graph for Example 1. Use it to specify a choice of D, m, and E that will in fact give a Lipschitz bow tie, showing that $\operatorname{Lip\,lim}_{x\to-1} x^2$ is in fact 1.

Theorem. *If* $\operatorname*{Lip\,lim}_{x\to a} f(x)=\ell$, *then* $\lim_{x\to a} f(x)=\ell$.

Proof. Let $\varepsilon>0$. By definition of the Lipschitz limit, there exist $D>0$, $E>0$ such that

$$|x-a|<D \text{ AND } x\neq a \implies |f(x)-\ell| < \left(\tfrac{E}{D}\right)|x-a|. \tag{1}$$

Case 1: $\varepsilon \le E$. In this case, pick

$$\delta = \left(\tfrac{D}{E}\right)\varepsilon.$$

If $|x-a|<\delta$ and $x\neq a$, then

$$\begin{aligned} |x-a| &< \frac{D}{E}\cdot\varepsilon \\ &\le \frac{D}{E}\cdot E \\ &= D. \end{aligned}$$

Since $|x-a|<D$ and $x \neq a$, implication (1) implies that

$$\begin{aligned} |f(x)-\ell| &< \left(\tfrac{E}{D}\right) \cdot |x-a| \\ &< \left(\tfrac{E}{D}\right) \cdot \delta \\ &= \left(\tfrac{E}{D}\right) \cdot \left(\tfrac{D}{E}\right) \varepsilon \\ &= \varepsilon. \end{aligned}$$

Case 2: $\varepsilon > E$. In this case, pick

$$\delta = D.$$

If $|x-a|<\delta$ and $x \neq a$, then $|x-a|<D$. Implication (1) then implies

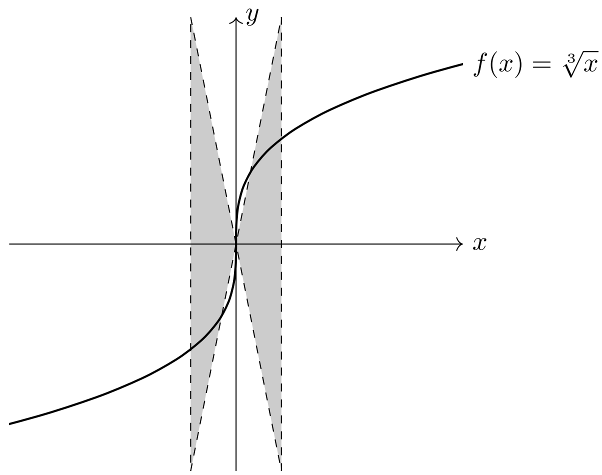
$$\begin{aligned} |f(x)-\ell| &< \left(\tfrac{E}{D}\right) \cdot |x-a| \\ &< \left(\tfrac{E}{D}\right) \cdot D \\ &= E \\ &< \varepsilon. \end{aligned}$$

□

Warning. *It is* not true *that*

$$\xcancel{\lim_{x\to a} f(x)=\ell \implies \operatorname{Lip}\lim_{x\to a} f(x)=\ell}$$

Example 3. For instance, as $x \to 0$, $\sqrt[3]{x}$ approaches the *limit* 0, but the cube root function has no *Lipschitz* limit as $x \to 0$.



No matter how steep and narrow we make the bow tie, it will never contain the function near 0. [Note: this is not a proof.]

MATH 131, LECTURE 12

INSTRUCTOR: CHARLES STAATS

1. Ways limits can fail to exist

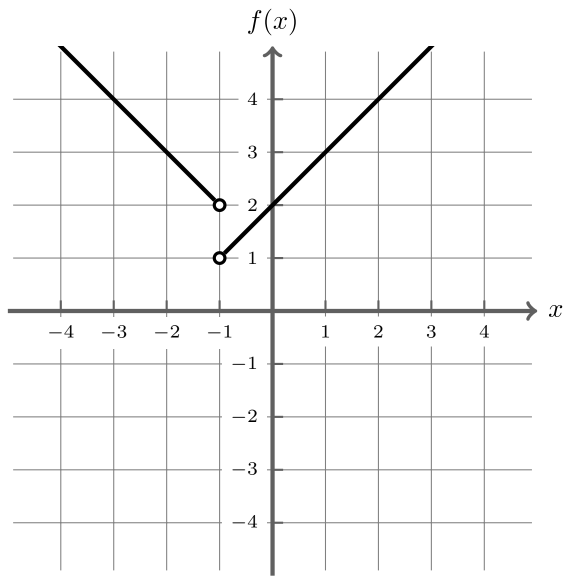
1.1. Jumps; one-sided limits. We consider the function f defined by

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

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

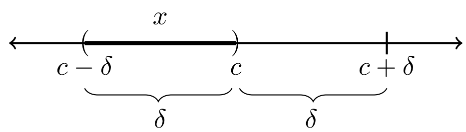
$$\lim_{x \to c^-} f(x) = \ell,$$

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":

Date: 29 October 2012.

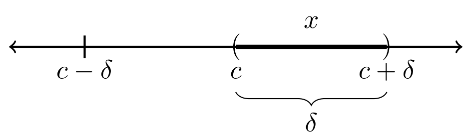


We say that "$f(x) \to \ell$ as $x \to c$ from the right," written

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

if

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



Exercise 1. *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 1. If

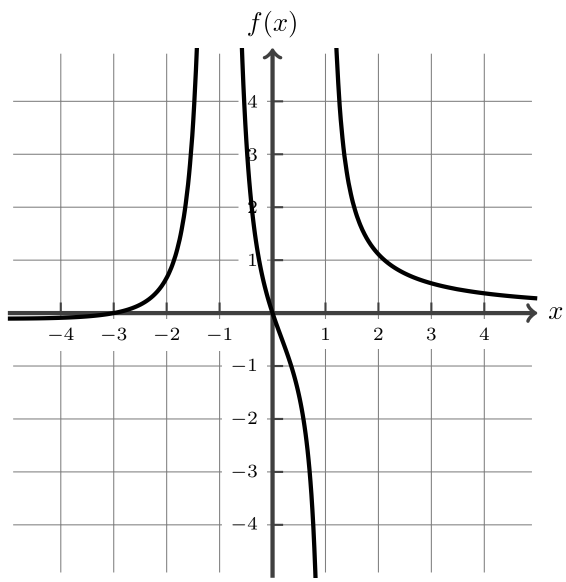
$$f(x) = \frac{1}{(x+1)^2} + \frac{1}{x-1},$$

then

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

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

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



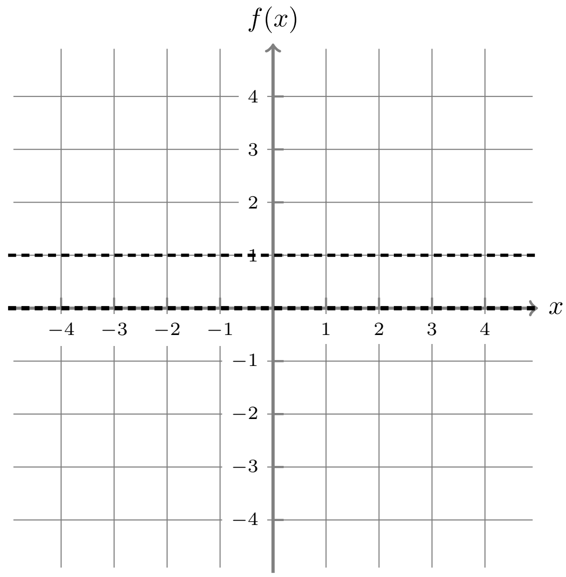
1.3. **Limits that just plain don't exist.** Consider the function g defined by

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

Consider

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

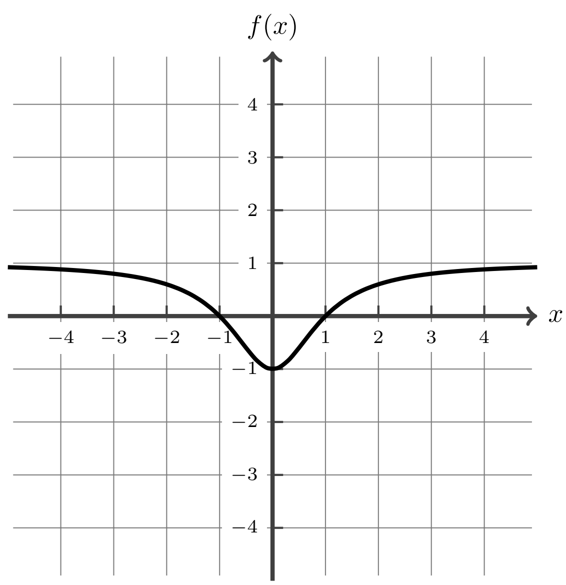
No matter how small x is, there are always smaller values at which f is 0 (say, $x = \frac{1}{n}$, for some really big integer n) and smaller values at which f is 1 (say, $x = \frac{1}{n}\sqrt{2}$, for an even bigger integer n). So, no version of the limit as $x \to 0$ can exist—not the left-hand limit,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



2. Limits as $x \to \infty$

Example 2. Consider the function f defined by

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



As x gets larger and larger, $f(x)$ gets closer and closer to 1. However, it makes no sense to say $f(\infty) = 1$, since ∞ is not a number (and even if it were, f is not defined there). Thus, we instead use the notion of the "limit as $x \to \infty$": we say that

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

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

Very informal: As x gets large, $f(x)$ gets close to ℓ.

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

Assignment 10 (due Wednesday, 31 October, a.k.a. Halloween)

Use the Intermediate Value Theorem to prove that, no matter what Diophantus of Alexandria[1] might have thought, $\sqrt{2}$ does, in fact, exist. (In other words, there exists a positive real number x_0 such that $x_0^2 = 2$; See Example 4 in Lecture 10.) This problem will be graded carefully.

Section 1.5, problems 29 and 30. Problem 30 will be graded carefully.

Section 1.6, problems 1, 3, and 5. Graph the function before answering the question. Problems 3 will be graded carefully.

Assignment 11 (due Friday, 2 November)

I have not yet chosen all the problems for this assignment. However, it will include the following problem, which requires some thought and thus should be started early:

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."]

[1] See Lecture 3, pp. 2 and 3.

MATH 131, LECTURE 13

INSTRUCTOR: CHARLES STAATS

1. Computing limits as $x \to \infty$

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.

(a) $$\lim_{x\to\infty} k = k$$

(b) $$\lim_{x\to\infty} \frac{1}{x} = 0$$

(c) $$\lim_{x\to\infty} [f(x) + g(x)] = \left[\lim_{x\to\infty} f(x)\right] + \left[\lim_{x\to\infty} g(x)\right]$$

(d) $$\lim_{x\to\infty} [f(x) - g(x)] = \left[\lim_{x\to\infty} f(x)\right] - \left[\lim_{x\to\infty} g(x)\right]$$

(e) $$\lim_{x\to\infty} [f(x) \cdot g(x)] = \left[\lim_{x\to\infty} f(x)\right] \cdot \left[\lim_{x\to\infty} g(x)\right]$$

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

(g) $$\lim_{x\to\infty} [f(x)]^n = \left[\lim_{x\to\infty} f(x)\right]^n$$

"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.*

Date: 31 October 2012.

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 1. (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 f, 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.

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

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

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

[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.

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 $\lim_{x\to\infty} k = k$.

2. 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.

3. 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

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

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.

Assignment 11 (due Friday, 2 November)

Section 1.5 problems 1–4, 27, and 28. Use the Main Limit Theorem, together with some algebra, to compute the limits. Problems 2, 4, and 28 will be graded carefully.

Section 1.6, Problems 7 and 8. Problem 8 will be graded carefully.

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

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

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."]

MATH 131, LECTURE 14

INSTRUCTOR: CHARLES STAATS

1. 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

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

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.

Date: 2 November 2012.

2. Thoughts

Consider how a speedometer works. One can imagine that it shows "instantaneous velocity," i.e., the speed of your car at each instance. However, in practice, the way it actually works is (I assume) that devices in the car measure the distance Δx that the car moves in a small amount of time Δt, and then reports the "instantaneous velocity" as

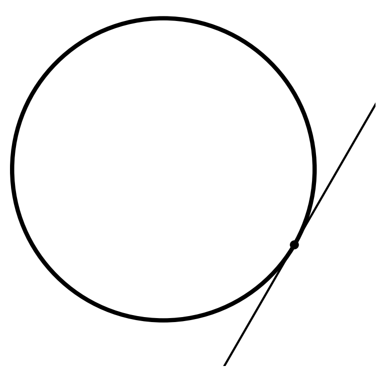
$$\frac{\Delta x}{\Delta t}.$$

When calculus was first invented, the inventors (Newton and Leibniz) did not really think about the derivative as a limit $\lim_{\Delta t \to 0} \Delta x/\Delta t$. Instead, they envisioned making Δt and Δx really, really small. The difficulty with this is that how small is "small enough" varies a lot. For a car traveling on the highway on cruise control, taking Δt to be one second will probably give you a pretty accurate value for the instantaneous velocity. But if you want to study the velocity of a car while it crashes into a concrete wall, I'm guessing the car stops so quickly that $\Delta t = 0.1$ seconds is not nearly small enough to get a good value.

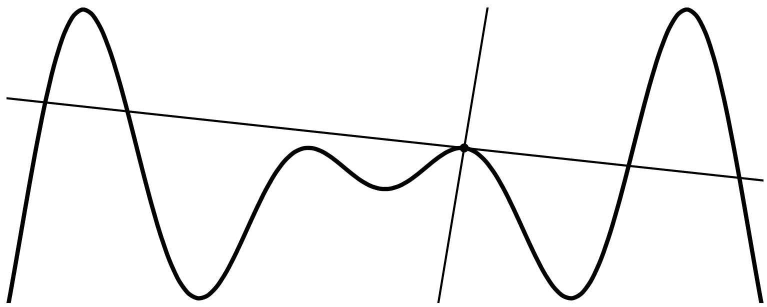
Newton and Leibniz tried to deal with this by saying that Δt and Δx should be "infinitesimal," that is, "infinitely small." Leibniz called these infinitesimal quantities dt and dx, and defined the derivative as their quotient dx/dt. The notion of numbers that are "infinitesimal but nonzero" does not really work logically. Nevertheless, it is a useful enough notion that physicists and chemists often still think about calculus this way.

3. The derivative as the slope of the tangent line

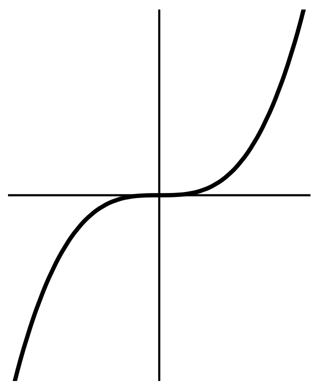
In classical geometry, the tangent to a curve was the line that somehow "touched the curve without crossing it." Euclid attempted to make this precise by describing the tangent as the line that intersected the curve in only one point. His definition works quite well for circles (and also ellipses, parabolas,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

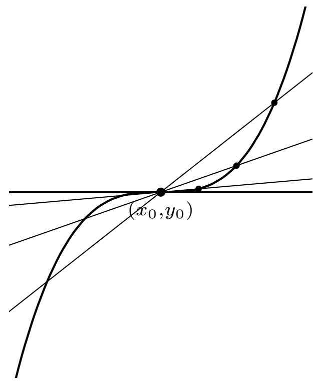
To remedy the situation, we consider another way of specifying a line: a point (x_0, y_0) together with a slope $\Delta y/\Delta x$. Thus, the secant line through (x_0, y_0) and (x, y) is the line passing through (x_0, y_0) with slope equal to

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

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 1. 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.

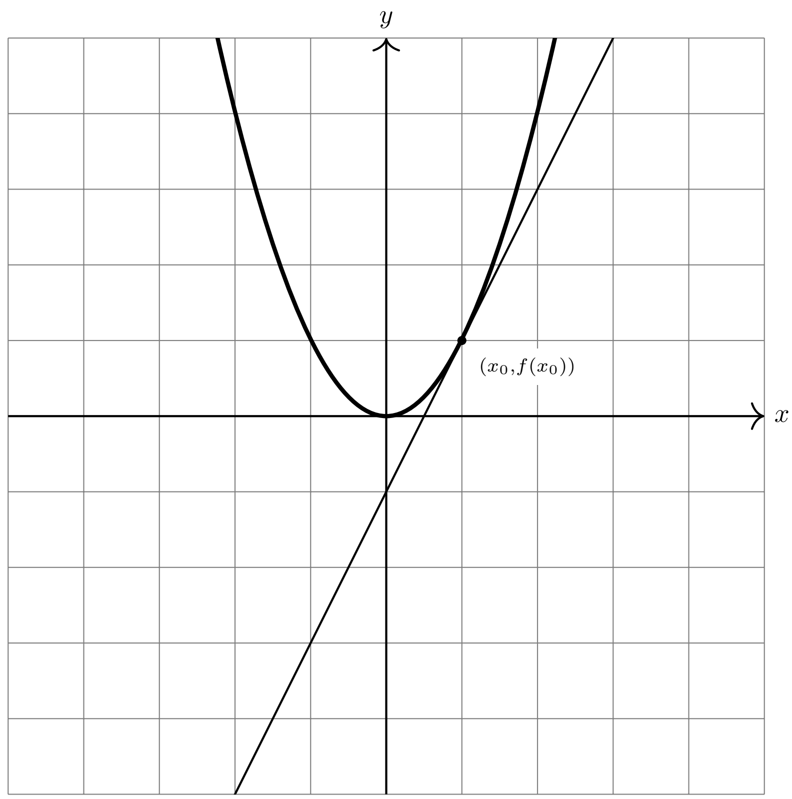
Solution. First, let's solve for Δy in terms of Δx:

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

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 $f'(x_0) = 2$:



□

4. 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—or, perhaps, "so close it might as well be equal." What this means, in this example, is that dx is "so small" that dx^2 can be treated as being equal to zero, 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 12 (due 5 November, 2012

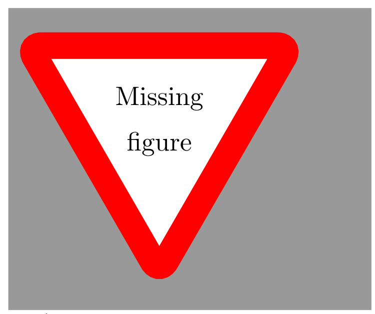
Section 2.2, Problems 45-48 and 51, 52. Be sure to follow the instructions carefully on 51 and 52; these problems are as much about *how* you find the derivative, as what answer you get. Problems 46, 48, and 52 will be graded carefully.

MATH 131, LECTURE 15

INSTRUCTOR: CHARLES STAATS

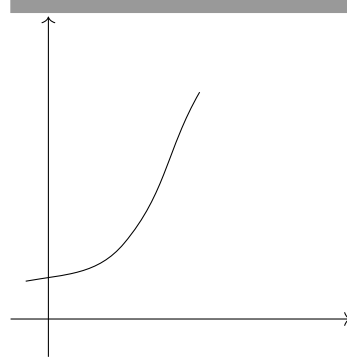
1. Computing the derivative: an example

First, recall the basic relations among the different quantities we are throwing around:



$$\begin{aligned} \Delta x &= x - x_0 & x &= x_0 + \Delta x \\ \Delta y &= y - y_0 & y &= y_0 + \Delta y \end{aligned}$$

$$\begin{aligned} y &= f(x) \\ &= f(x_0 + \Delta x) \\ \Delta y &= y - y_0 \\ &= f(x) - f(x_0) \\ &= f(x_0 + \Delta x) - f(x_0) \end{aligned}$$



Example 1. Let the function f be defined by

$$f(x) = \tfrac{1}{2}x^3.$$

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

Date: 5 November 2012.

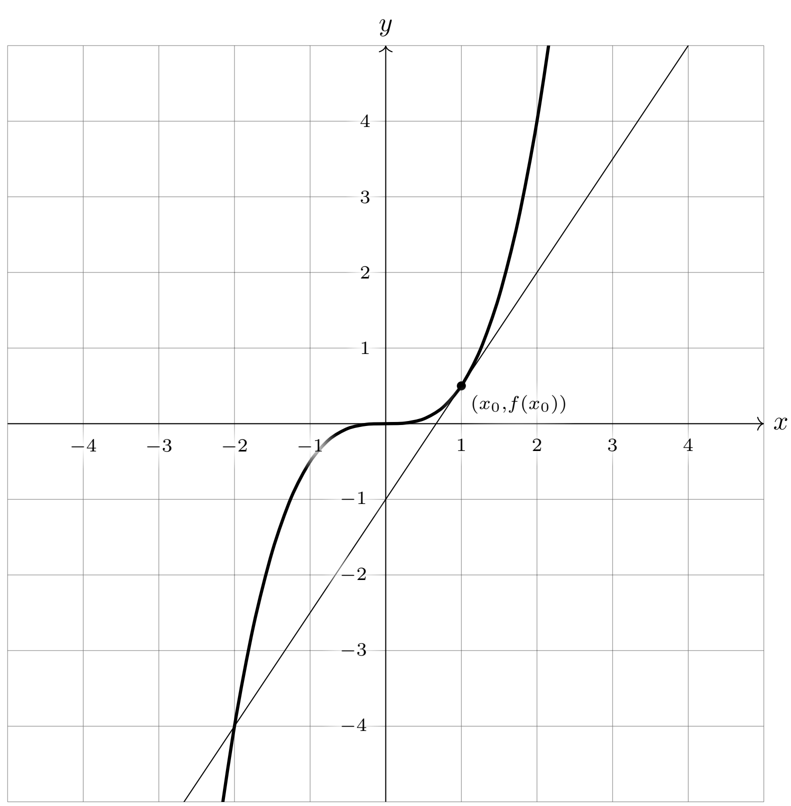
Solution. First, let's solve for Δy in terms of Δx:

$$\begin{aligned}\Delta y &= f(x_0+\Delta x) - f(x_0)\\ &= \tfrac{1}{2}(1+\Delta x)^3 - \tfrac{1}{2}(1)^3\\ &= \tfrac{1}{2}\Big((1+\Delta x)^3 - 1^3\Big)\\ &= \tfrac{1}{2}\Big(\big(1+3\Delta x+3(\Delta x)^2+(\Delta x)^3\big) - 1\Big)\\ &= \tfrac{1}{2}\Big(3\Delta x+3(\Delta x)^2+(\Delta x)^3\Big)\\ &= \tfrac{3}{2}\Delta x+\tfrac{3}{2}(\Delta x)^2+\tfrac{1}{2}(\Delta x)^3.\end{aligned}$$

Thus, we have

$$\begin{aligned}f'(x_0) &= \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{\frac{3}{2}\Delta x+\frac{3}{2}(\Delta x)^2+\frac{1}{2}(\Delta x)^3}{\Delta x}\\ &= \lim_{\Delta x\to 0}\tfrac{3}{2}+\tfrac{3}{2}\Delta x+\tfrac{1}{2}(\Delta x)^2\\ &= \tfrac{3}{2}.\end{aligned}$$

Now, we plot the function $y = f(x)$, together with line passing through $(x_0, f(x_0)) = (1, \frac{1}{2})$ and having slope $f'(x_0) = \frac{3}{2}$:



□

2. Infinitesimals

Consider how a speedometer works. One can imagine that it shows "instantaneous velocity," i.e., the speed of your car at each instance. However, in practice, the way it actually works is (I assume) that devices in the car measure the distance Δx that the car moves in a small amount of time Δt, and then reports the "instantaneous velocity" as

$$\frac{\Delta x}{\Delta t}.$$

When calculus was first invented, the inventors (Newton and Leibniz) did not really think about the derivative as a limit $\lim_{\Delta t \to 0} \Delta x/\Delta t$. Instead, they envisioned making Δt and Δx really, really small. The difficulty with this is that how small is "small enough" varies a lot. For a car traveling on the highway on cruise control, taking Δt to be one second will probably give you a pretty accurate value for the instantaneous velocity. But if you want to study the velocity of a car while it crashes into a concrete wall, I'm guessing the car stops so quickly that $\Delta t = 0.1$ seconds is not nearly small enough to get a good value.

Newton and Leibniz tried to deal with this by saying that Δt and Δx should be "infinitesimal," that is, "infinitely small." Leibniz called these infinitesimal quantities dt and dx, and defined the derivative as their quotient dx/dt. The notion of numbers that are "infinitesimal but nonzero" does not really work logically. Nevertheless, it is a useful enough notion that physicists and chemists often still think about calculus this way.

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—or, perhaps, "so close it might as well be equal." What this means, more or less, is that you can do the entire computation treating dx and dy as ordinary numbers, including dividing by dx at the end; once you have something of the form

$$\frac{dy}{dx} =$$

you can treat dx as being equal to zero in the expression on the right. This does not really make rigorous sense—dx is either equal to zero, or it isn't; it can't somehow become zero after you've finished your computation—but it usually gives the right answer once you get used to it, and it can make computations quicker and more intuitive (assuming you know when you can, and cannot, set dx equal to zero...).

Example 2. Let the function f be defined by

$$f(x) = \tfrac{1}{2}x^3.$$

Use infinitesimals to compute the derivative of f at $x = 1$.

Solution. To start with, we treat the "infinitesimal changes" dx and dy exactly as though they were more conventional changes Δx and Δy.

$$\begin{aligned} dy &= f(x+dx) - f(x) \\ &= \tfrac{1}{2}(1+dx)^3 - \tfrac{1}{2}(1)^3 \\ &= \tfrac{1}{2}\left[(1+dx)^3 - 1^3\right] \\ &= \tfrac{1}{2}\left[(1+3dx+3dx^2+dx^3) - 1\right] \\ &= \tfrac{3}{2}dx + \tfrac{3}{2}dx^2 + \tfrac{1}{2}dx^3 \\ \frac{dy}{dx} &= \tfrac{3}{2} + \cancelto{0}{\tfrac{3}{2}dx} + \cancelto{0}{\tfrac{1}{2}dx^2} \end{aligned}$$

when evaluated at the point $x_0 = 1$. □

What often happens is that when you get to the end of the computation, if you see an expression like $4 + dx$ or $dx + 3dx^2$, you can "ignore" dx (in the first case) or $3dx^2$ (in the second case), because it is "negligibly small" compared to the other summand.

Exercise 3. *Use infinitesimals to find the derivative of the function f defined by*

$$f(x) = \frac{1}{x}$$

at the point $x_0 = 1$.

Assignment 13 (due Wednesday, 7 November, 2012)

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 3 will be graded carefully.

Section 2.2, Problems 5 and 7. Neither of these will 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 $y = f(x)$. This problem will be graded carefully.

MATH 131, LECTURE 16

INSTRUCTOR: CHARLES STAATS

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., $f'(x_0) = g'(x_0)$.

Date: 7 November 2012.

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."

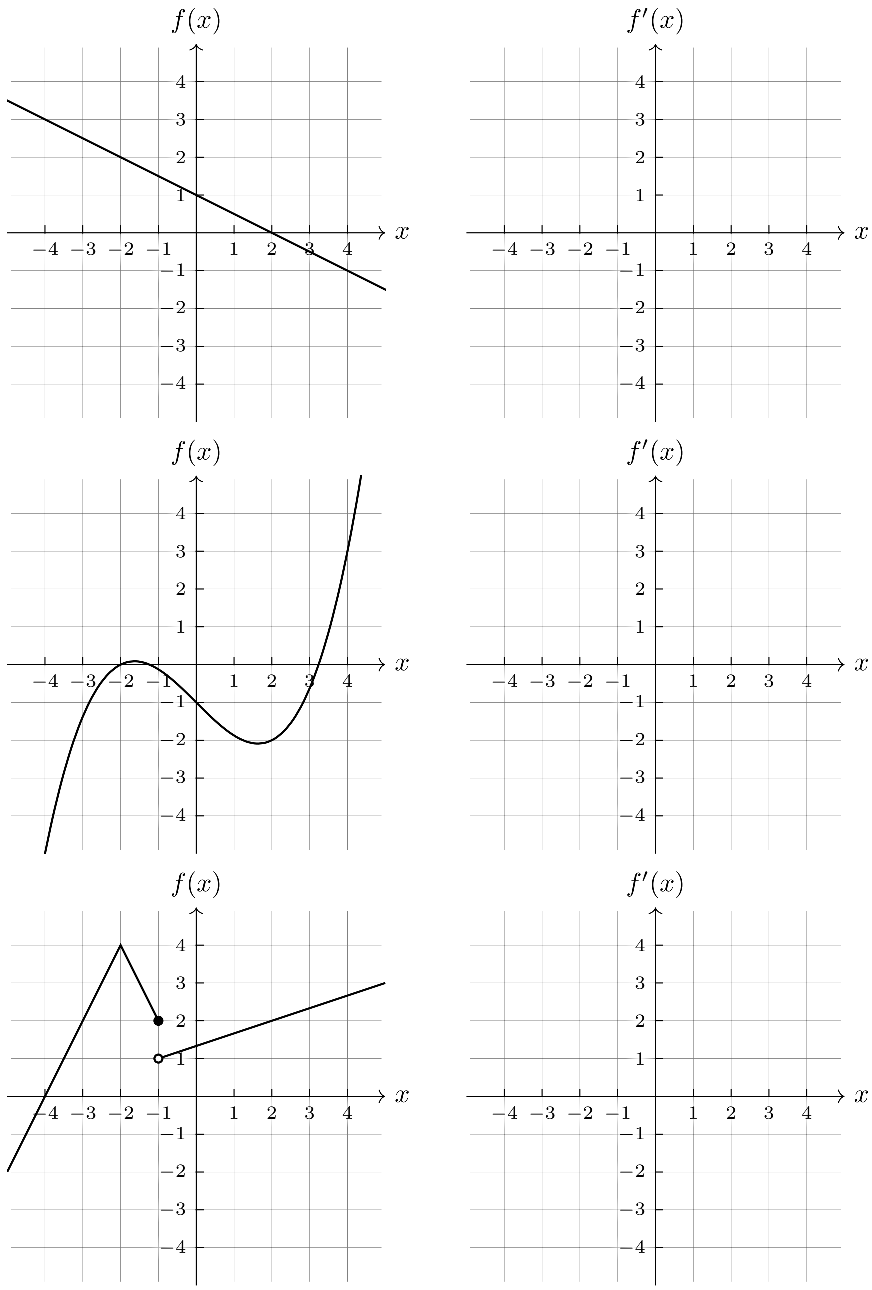
On the other hand, some of you may be thinking that it was obvious that the derivative is a function. You may even feel a bit smug about the fact that this "revelation" was clear to you from the beginning. Perhaps you should. But I think it is more likely that you were not following my lectures closely, 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 1. 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.



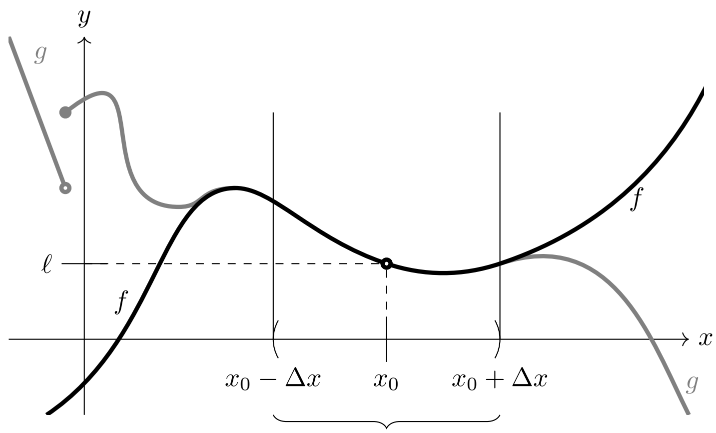
3. Local nature of the limit (and derivative)

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 $(x_0 - \Delta x, x_0 + \Delta x)$.



$f = g$ on this interval, so
$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x)$$

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, 9 November, 2012)

Section 2.2, Problems 6, 8, and 37–44. Problems 6, 38, 40, 42, and 44 will be graded carefully.

The following problem is not due until Monday. However, it is a difficult problem and the only tutorial between now and Monday is tomorrow; so I strongly suggest you start on it now.

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 $\frac{dy}{dx}\Delta x$."

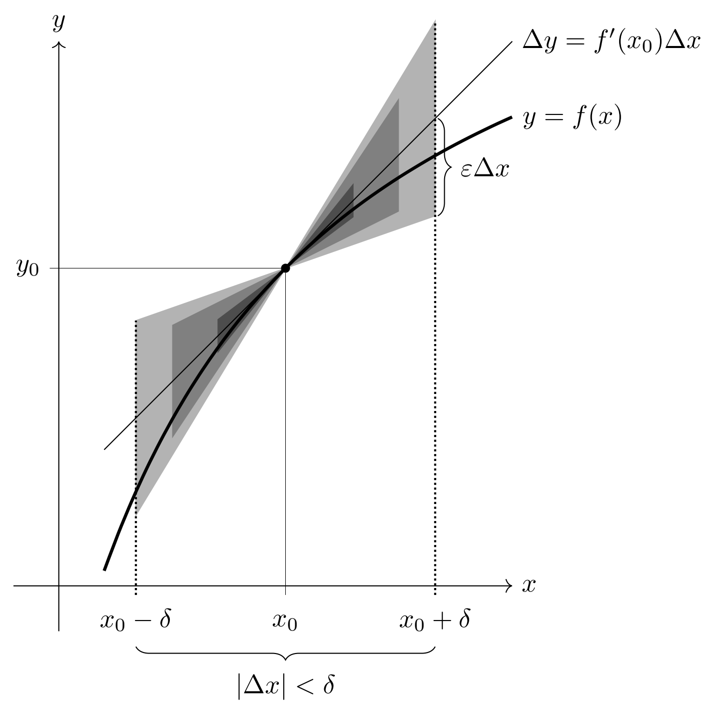


Figure 1. When Δx is small, then Δy is approximated by $f'(x_0)\Delta x$. In other words, for Δx small, the function is approximated by its tangent line (which is defined by $\Delta y = f'(x_0)\Delta x$). More precisely, the function is contained in a narrow cone about the tangent line. The width of 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.

MATH 131, LECTURE 17

INSTRUCTOR: CHARLES STAATS

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 1. 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 2. 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 3. Let f be the function defined by $f(x) = x^2 + x - 3$. Compute a formula for the function f'.

Date: 9 November 2012.

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 4. 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*

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

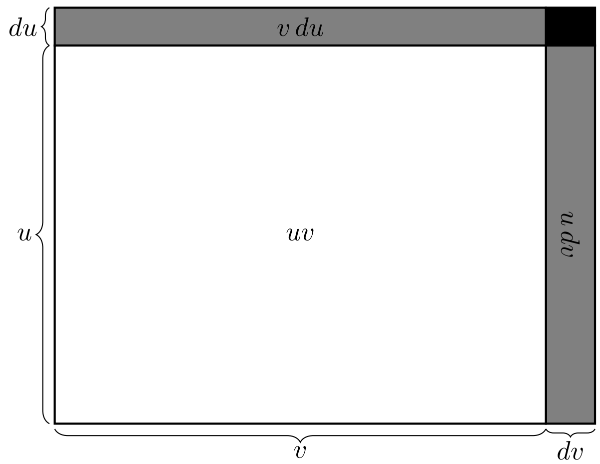


FIGURE 1. A visual illustration of the product rule. The area of the white rectangle is uv; the area of the total rectangle is $(u + du)(v + dv)$; and the change in area, $d(uv)$, is their difference. The black rectangle, with area $du\,dv$, is so small that its contribution "can be neglected."

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 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

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

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 5. 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, 12 November, 2012)

Section 2.2, Problems 11, 12, 23, 24, 53, and 54. Follow the instructions. The even-numbered problems will be graded carefully.

The following problem 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 $\frac{dy}{dx}\Delta x$."

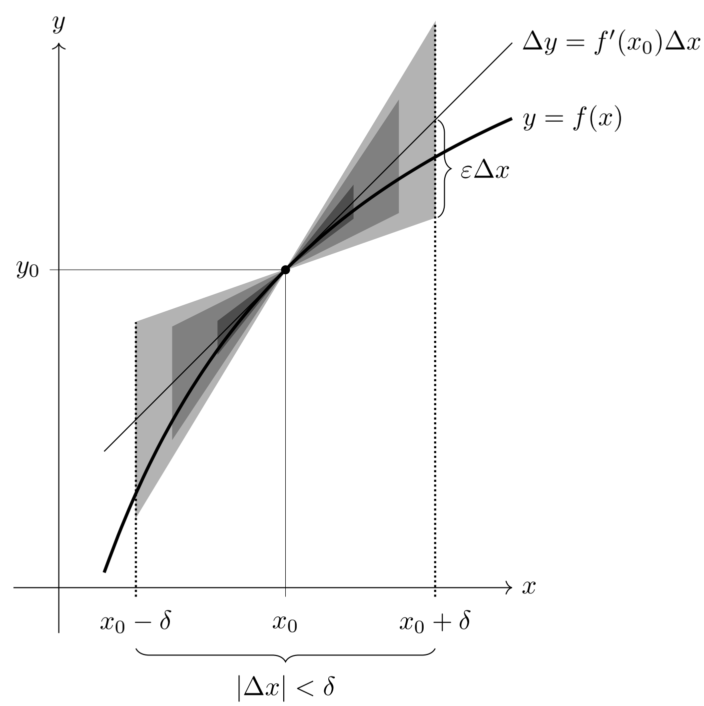


Figure 2. When Δx is small, then Δy is approximated by $f'(x_0)\Delta x$. In other words, for Δx small, the function is approximated by its tangent line (which is defined by $\Delta y = f'(x_0)\Delta x$). More precisely, the function is contained in a narrow cone about the tangent line. The width of 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.

MATH 131, LECTURE 18: RULES FOR DIFFERENTIATION

INSTRUCTOR: CHARLES STAATS

1. Test Monday, 19 November

The test will cover Lectures 1–19, with emphasis on Lectures 10–19; and Assignments 1–17, with emphasis on Assignments 9–17. (Reference note: Lecture 19 will be given on Wednesday; Assignment 17 is due Friday.) The test will focus on derivatives, but will likely include a bit about continuity and/or infinite limits and/or limits as $x \to \pm\infty$.

You should not assume that if you did well on the last test, you can easily do well on this test. As a general rule, students do worse on the second test in Math 131 than they did on the first test. Thus, I suggest you start studying yesterday, if not before. 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.

2. Introduction

The process of finding the derivative of a function, often called *differentiating* the function, 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.

3. The product rule

Recall the product rule, which I introduced last lecture:

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

Personally, I find it easier to remember the infinitesimal version:

$$d(uv) = u\,dv + v\,du.$$

Finally, some of you may find an alternative version easier to apply:

$$\begin{aligned}\frac{d}{dx}\big(f(x)\cdot g(x)\big) &= f(x)g'(x) + g(x)f'(x)\\ &= f(x)g'(x) + f'(x)g(x).\end{aligned}$$

Example 1. Use the product rule to find (in this order) the derivatives of x^2, x^3, and x^4 with respect to x.

Date: 12 November 2012.

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

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

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

□

4. 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

$$(f+g)' = f'+g'.$$ □

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 time, we're going to derive it from the product rule.

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

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

Proof.

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

5. 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 2. Differentiate $x^2 - 4x + 1$.

Solution.

$$\begin{aligned}
\frac{d}{dx}(x^2 - 4x + 1) &= \frac{d}{dx}(x^2) - \frac{d}{dx}(4x) + \frac{d}{dx}(1) && \text{(sum rule)} \\
&= \frac{d}{dx}(x^2) - 4\frac{d}{dx}(x) + 0 && \text{(constant multiple; constant)} \\
&= 2x - 4 \cdot 1 + 0 && \text{(power rule)} \\
&= 2x - 4.
\end{aligned}$$

□

Example 3. Differentiate $2x^3 - \frac{1}{2}x^2 - x + \frac{17246}{937}$.

Solution.

$$\begin{aligned}
\frac{d}{dx}\left(2x^3 - \tfrac{1}{2}x^2 - x + \tfrac{17246}{937}\right) &= 2 \cdot 3x^2 - \tfrac{1}{2} \cdot 2x - 1 + 0 \\
&= 6x^2 - x - 1.
\end{aligned}$$

□

6. 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:

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

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

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

$$\left.\frac{d}{dx}(uv)\right|_{x=x_0} = u_0 \left.\frac{dv}{dx}\right|_{x=x_0} + v_0 \left.\frac{du}{dx}\right|_{x=x_0}.$$

Proof. First, we write $\Delta(uv)$ in terms of Δu and Δv:

$$\begin{aligned}
\Delta(uv) &= uv - u_0 v_0 \\
&= (u_0 + \Delta u)(v_0 + \Delta v) - u_0 v_0 \\
&= u_0 v_u + u_0 \Delta v + v_0 \Delta u + \Delta u \Delta v - u_0 v_0 \\
&= u_0 \Delta v + v_0 \Delta u + \Delta u \Delta v.
\end{aligned}$$

[Notice how closely this resembles the infinitesimal version.] Now, we apply the definition[1] of the derivative as a limit:

$$\begin{aligned}
\frac{d}{dx}(uv)\bigg|_{x=x_0} &= \lim_{\Delta x\to 0}\frac{\Delta(uv)}{\Delta x} \\
&= \lim_{\Delta x\to 0}\frac{u_0\Delta v + v_0\Delta u + \Delta u\Delta v}{\Delta x} \\
&= \lim_{\Delta x\to 0} u_0\frac{\Delta v}{\Delta x} + v_0\frac{\Delta u}{\Delta x} + \frac{\Delta u}{\Delta x}\cdot\frac{\Delta v}{\Delta x}\cdot\Delta x \\
&= u_0\left.\frac{dv}{dx}\right|_{x=x_0} + v_0\left.\frac{du}{dx}\right|_{x=x_0} + \left(\left.\frac{du}{dx}\right|_{x=x_0}\right)\left(\left.\frac{dv}{dx}\right|_{x=x_0}\right)\cdot 0 \\
&= u_0\left.\frac{dv}{dx}\right|_{x=x_0} + v_0\left.\frac{du}{dx}\right|_{x=x_0}.
\end{aligned}$$

□

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 $du\,dv = 0$.

7. 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

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

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 4. Differentiate $(x+1)^{500}$.

Solution.

$$\begin{aligned}
\frac{d}{dx}(x+1)^{500} &= 500(x+1)^{499}\cdot\frac{d}{dx}(x+1) \\
&= 500(x+1)^{499}\cdot 1 \\
&= 500(x+1)^{499}.
\end{aligned}$$

□

[1] Or rather, one of the equivalent definitions.

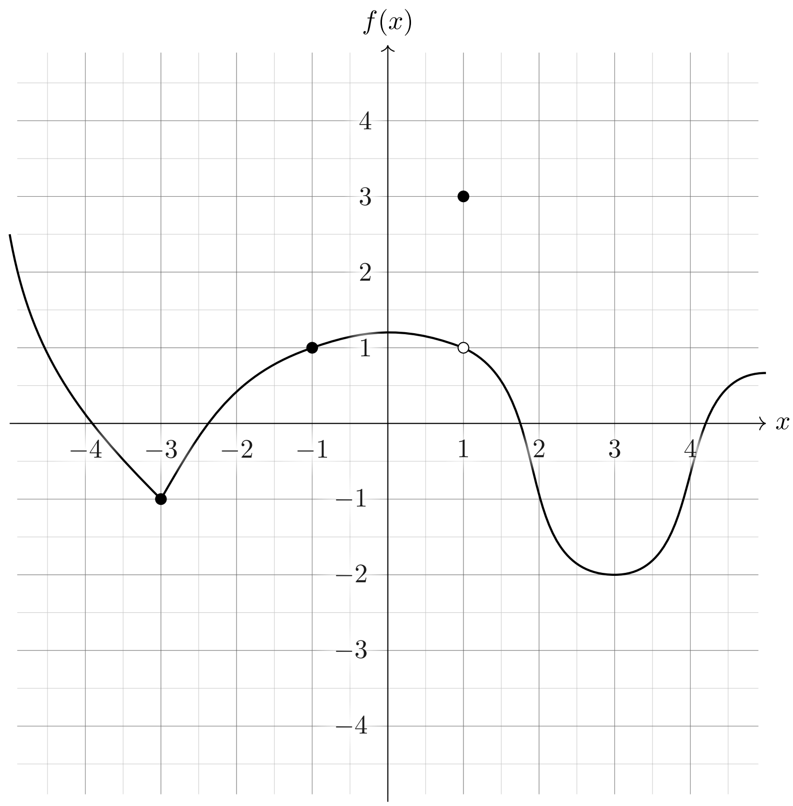
It would have been possible, but very hard, to differentiate this by expanding out all 501 terms of the polynomial and then applying the techniques of the first section.

Assignment 16 (due Wednesday, 14 November)

Section 2.2, Problems 13. Follow the instructions. This problem will be graded carefully.

Section 2.3, Problems 1–4, 11–12, and 23–26. (Hint: 23–26 are easier if you use the product rule.) The even-numbered problems will be graded carefully.

Sketch the graph of $f'(x)$ for the function f graphed below. Hint: First, make a table of the values of $f'(x)$ for some specific values of x.



Assignment 17 (due Friday, 16 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 4 at the end of Lecture 18). Problems 2 and 4 will be graded carefully.

MATH 131, LECTURE 19

INSTRUCTOR: CHARLES STAATS

1. Test Monday, 19 November

The test will cover Lectures 1–19, with emphasis on Lectures 10–19; and Assignments 1–17, with emphasis on Assignments 9–17. (Reference note: Lecture 19 is today's lecture; Assignment 17 is due Friday.) The test will focus on derivatives, but will likely include a bit about continuity and/or infinite limits and/or limits as $x \to \pm\infty$.

You should not assume that if you did well on the last test, you can easily do well on this test. As a general rule, students do worse on the second test in Math 131 than they did on the first test. Thus, I suggest you start studying yesterday, if not before. 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.

2. 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

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

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,

$$\frac{dy}{dx} = f'(g(x))g'(x).$$

Date: 14 November 2012.

Example 1. Differentiate $(x+1)^{500}$.

Solution.

$$\begin{aligned}\frac{d}{dx}(x+1)^{500} &= 500(x+1)^{499} \cdot \frac{d}{dx}(x+1)\\ &= 500(x+1)^{499} \cdot 1\\ &= 500(x+1)^{499}. \qquad \square\end{aligned}$$

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.

3. Differentiating quotients

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 2. 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.

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

If you want more examples of this sort of computation, you should review Lecture 17.

We can use the Chain Rule together with the Product Rule and Example 2 (page 2) to differentiate much more general quotients.

Example 3. Find $\frac{d}{dx}\left(\frac{1}{x-1}\right)$.

Recall, from Example 2, that $D_x(1/x) = -1/x^2$.

Solution.

$$\begin{aligned}\frac{d}{dx}\left(\frac{1}{x-1}\right) &= \frac{-1}{(x-1)^2} \cdot \frac{d}{dx}(x-1)\\ &= \frac{-1}{(x-1)^2}. \qquad \square\end{aligned}$$

Example 4. Let
$$f(x) = \frac{x+1}{x-1}.$$
Compute $f'(x)$.

Solution.
$$\begin{aligned}
\frac{d}{dx}\left(\frac{x+1}{x-1}\right) &= \frac{d}{dx}\left((x+1)\cdot\frac{1}{x-1}\right) \\
&= (x+1)\frac{d}{dx}\left(\frac{1}{x-1}\right) \\
&\quad + \frac{1}{x-1}\frac{d}{dx}(x+1) && \text{(Product Rule)} \\
&= (x+1)\cdot\frac{-1}{(x-1)^2}\cdot\frac{d}{dx}(x-1) \\
&\quad + \frac{1}{x-1}\cdot 1 && \text{(Chain Rule)} \\
&= -\frac{x+1}{(x-1)^2} + \frac{1}{x-1} \\
&= \frac{-(x+1)+(x-1)}{(x-1)^2} \\
&= \frac{-2}{(x-1)^2}.
\end{aligned}$$
□

4. 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:
$$\begin{aligned}
d(uv) &= (u+du)(v+dv) - uv \\
&= uv + u\,dv + v\,du + du\,dv - uv \\
&= u\,dv + v\,du + du\,dv
\end{aligned}$$
$$\begin{aligned}
\frac{d}{dx}(uv) &= u\frac{dv}{dx} + v\frac{du}{dx} + \frac{du}{dx}\cancelto{0}{dv} \\
&= u\frac{dv}{dx} + v\frac{du}{dx}.
\end{aligned}$$
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
$$\left.\frac{d}{dx}(uv)\right|_{x=x_0} = u_0\left.\frac{dv}{dx}\right|_{x=x_0} + v_0\left.\frac{du}{dx}\right|_{x=x_0}.$$

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:

$$\begin{aligned}\frac{d}{dx}(uv)\bigg|_{x=x_0} &= \lim_{\Delta x\to 0}\frac{\Delta(uv)}{\Delta x} \\ &= \lim_{\Delta x\to 0}\frac{u_0\Delta v + v_0\Delta u + \Delta u\Delta v}{\Delta x} \\ &= \lim_{\Delta x\to 0} u_0\frac{\Delta v}{\Delta x} + v_0\frac{\Delta u}{\Delta x} + \frac{\Delta u}{\Delta x}\cdot\frac{\Delta v}{\Delta x}\cdot\Delta x \\ &= u_0\,\frac{dv}{dx}\bigg|_{x=x_0} + v_0\,\frac{du}{dx}\bigg|_{x=x_0} + \left(\frac{du}{dx}\bigg|_{x=x_0}\right)\left(\frac{dv}{dx}\bigg|_{x=x_0}\right)\cdot 0 \\ &= u_0\,\frac{dv}{dx}\bigg|_{x=x_0} + v_0\,\frac{du}{dx}\bigg|_{x=x_0}.\end{aligned}$$

□

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 $\dfrac{du\,dv}{dx} = 0$.

[1] Or rather, one of the equivalent definitions.

Assignment 17 (due Friday, 16 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 1 at the end of Lecture 18). Problems 2 and 4 will be graded carefully.

MATH 131, LECTURE 20

INSTRUCTOR: CHARLES STAATS

1. Test Monday, 19 November

The test will cover Lectures 1–19, with emphasis on Lectures 10–19; and Assignments 1–17, with emphasis on Assignments 9–17. (Reference note: Lecture 19 was Wednesday's lecture; Assignment 17 was due today.) The test will focus on derivatives, but will likely include a bit about continuity and/or infinite limits and/or limits as $x \to \pm\infty$.

You should not assume that if you did well on the last test, you can easily do well on this test. As a general rule, students do worse on the second test in Math 131 than they did on the first test. Thus, I suggest you start studying yesterday, if not before. 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.

2. Proving the Chain Rule—the idea

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.

Date: 16 November 2012.

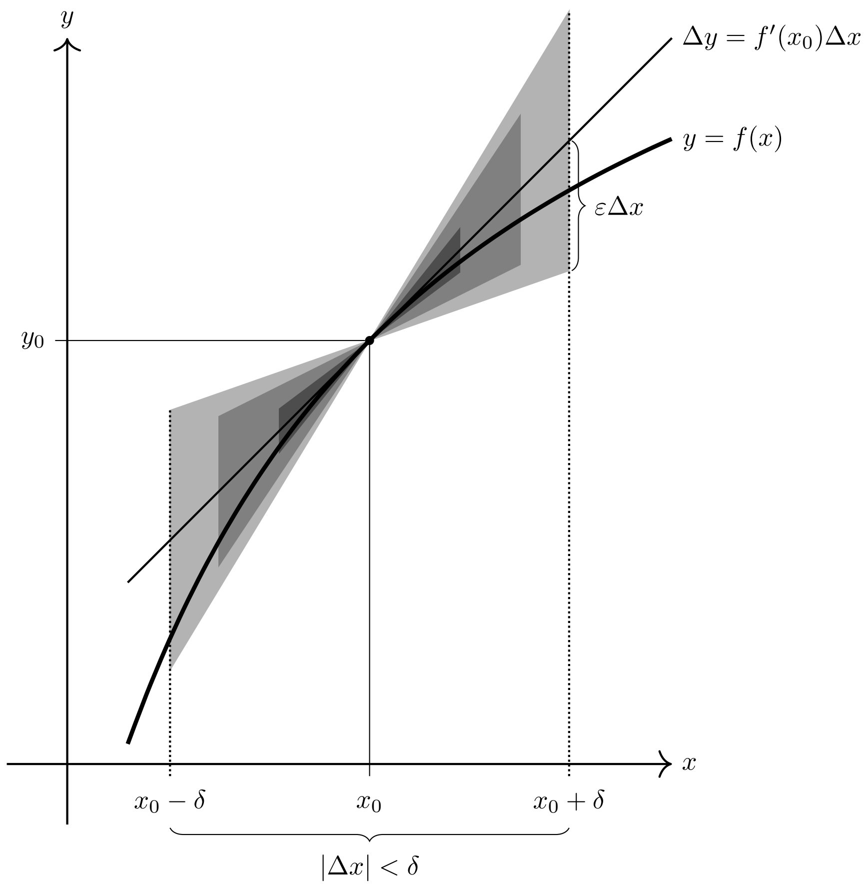


FIGURE 1. When Δx is small, then Δy is approximated by $f'(x_0)\Delta x$. In other words, for Δx small, the function is approximated by its tangent line (which is defined by $\Delta y = f'(x_0)\Delta x$). More precisely, the function is contained in a narrow cone about the tangent line. The width of 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.

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 $u = 2x + 1$. $\square$

Solution (short version).

$$\begin{aligned}\frac{d}{dx}(2x+1)^3 &= 3(2x+1)^2 \cdot \frac{d}{dx}(2x+1) \\ &= 3(2x+1)^2 \cdot 2 \\ &= 6(2x+1)^2.\end{aligned}$$ $\square$

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 Wednesday, 21 November)

There will be one, but I have not yet selected it.

MATH 131, LECTURE 21

INSTRUCTOR: CHARLES STAATS

1. 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 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,"

$$\frac{dy}{dx} = f'(g(x))g'(x).$$

Date: 21 November 2012.

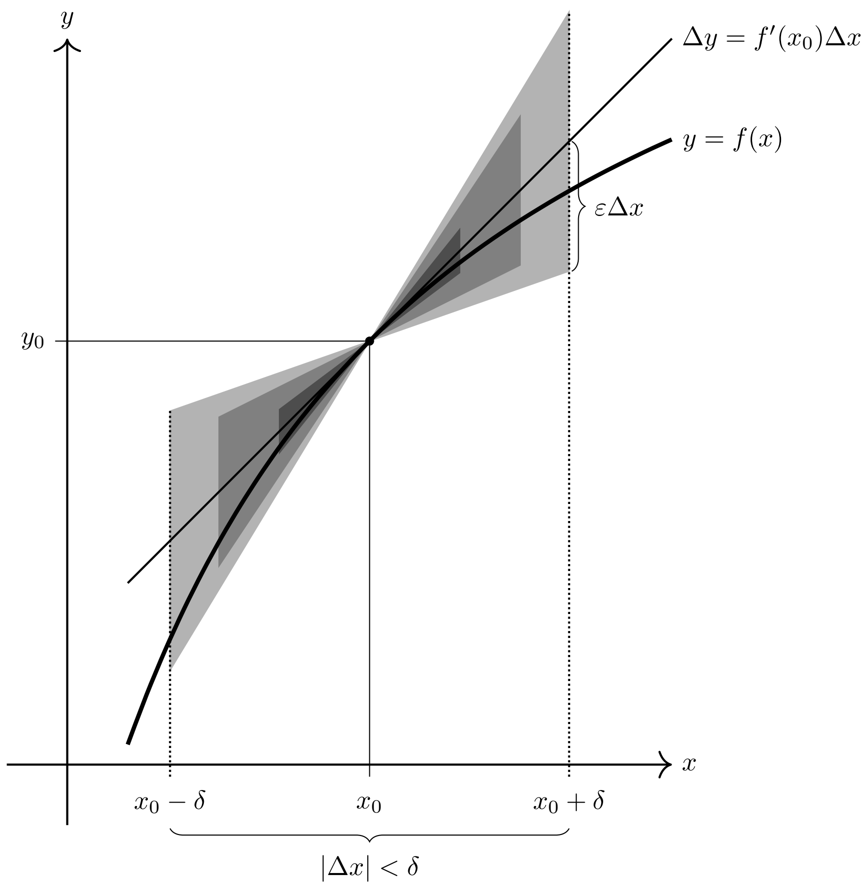


FIGURE 1. When Δx is small, then Δy is approximated by $f'(x_0)\Delta x$. In other words, for Δx small, the function is approximated by its tangent line (which is defined by $\Delta y = f'(x_0)\Delta x$). More precisely, the function is contained in a narrow cone about the tangent line. The width of 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.

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.

2. 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

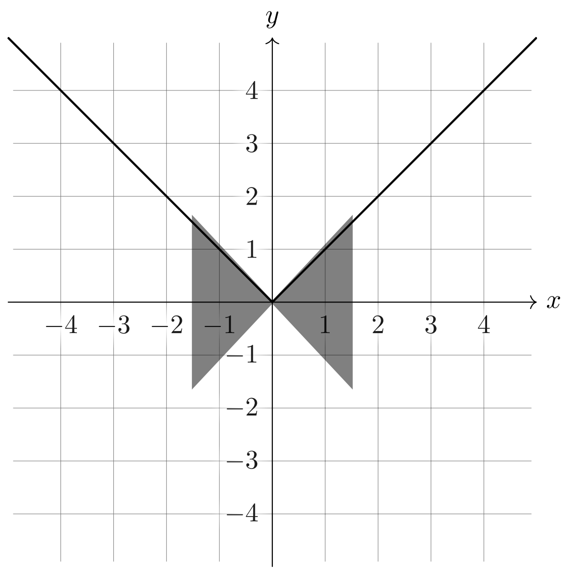
$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

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

$$\lim_{h\to 0^-} \frac{|0+h|-|0|}{h} = \lim_{h\to 0} \frac{-h}{h} = -1$$
$$\lim_{h\to 0^+} \frac{|0+h|-|0|}{h} = \lim_{h\to 0} \frac{h}{h} = 1.$$

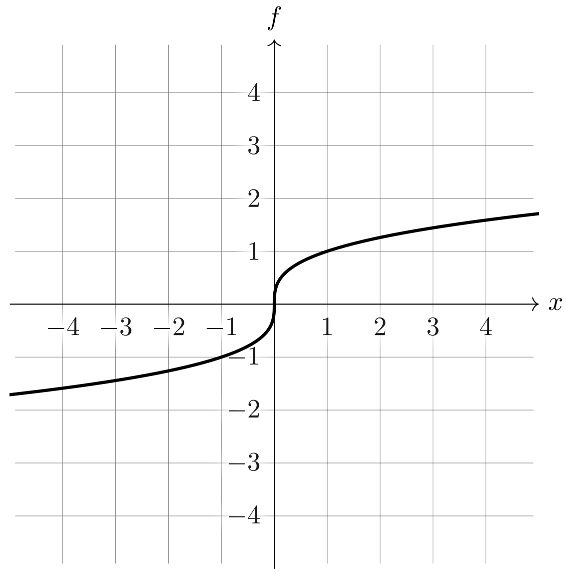
Since the one-sided limits are not equal, the two-sided limit

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

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

$$f'(0) = \lim_{h \to 0} \frac{(0+h)^{1/3} - 0^{1/3}}{h},$$

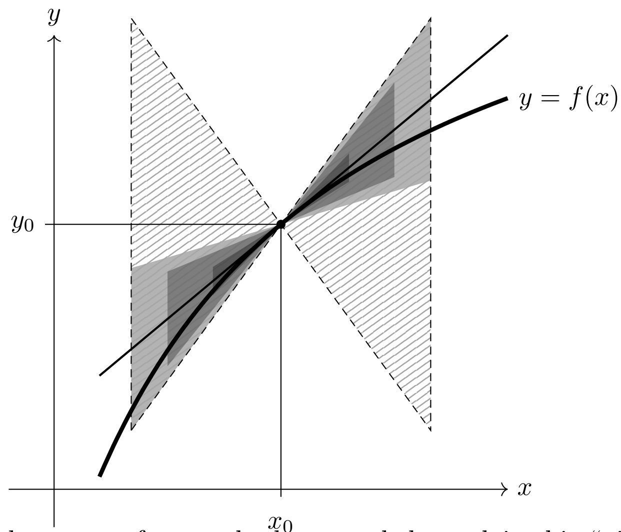
one will find that the limit is ∞. Since this derivative is not *finite*, we still say that f is not differentiable at 0.

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 without words:"



In case you are lost, here are a few words that may help explain this "picture proof": If f is differentiable at x_0, then we can draw a narrow cone centered at (x_0, y_0) containing f; we can then draw a Lipschitz bow tie containing this cone, and hence containing f. □

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.*

3. 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.

Assignment 19 (due Monday, 26 November)

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.

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

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 $\dfrac{d}{dx}\left(\dfrac{dy}{dx}\right)$ 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.

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

$$\frac{dx}{dt}$$

Date: Monday, 26 November 2011.

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 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

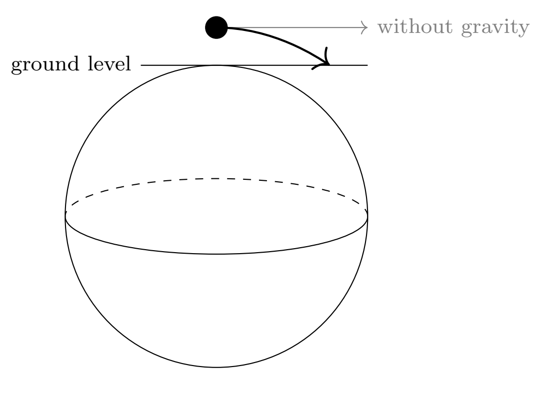


FIGURE 1. The spaceship is falling, but it's moving sidewise so quickly that by the time it reaches "ground level," it has not actually gotten any closer to the surface of the earth.

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 figure out 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:

$$\begin{aligned} y'(t) &= -10t + a \\ y''(t) &= -10. \end{aligned}$$

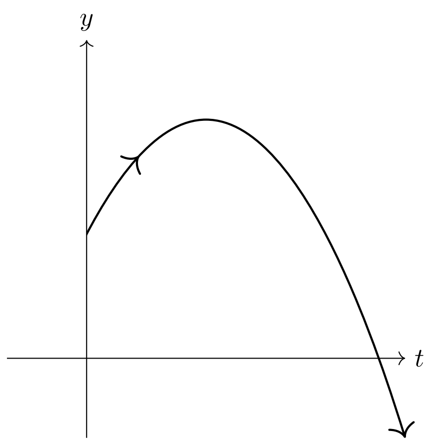
As it turns out, these are the only functions that satisfy this differential equation, although we will not see why until next quarter. Thus, any time you toss or drop an object, its vertical position is described by

$$y = -5t^2 + at + b$$

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 seems to scintillate with reflected light; cords of water seem to twist together, like the muscles in a Michelangelo drawing of an arm.

[1]If this were a physics course, we'd use the word "forces."



(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

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 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 Wednesday, 28 November)

Section 2.3, Problems 39–40. Problem 40 will be graded carefully.

Section 2.6, Problems 1–4 and 9–10. Problems 2, 4, and 10 will be graded carefully.

Assignment 21 (due Friday, 30 November)

Section 2.6, Problems 7–8 and 11–12. Problems 8 and 12 will be graded carefully.

Section 2.7, Problems 1–2. 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.)

MATH 131, LECTURE 23

CHARLES STAATS

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

$$f'(x) = \frac{1}{2\sqrt{x}}. \qquad \square$$

Date: Wednesday, 28 November 2011.

Exercise 2. *Use implicit differentiation to show that if $y = -\sqrt{x}$, then*

$$\frac{dy}{dx} = -\frac{1}{2\sqrt{x}}.$$

Solution. This time, y also satisfies the equation $y^2 = x$, since $\left(-\sqrt{x}\right)^2 = (-1)^2\left(\sqrt{x}\right)^2 = 1 \cdot x = x$. Thus, we can differentiate implicitly:

$$\begin{aligned} y^2 &= x \\ 2y \cdot \frac{dy}{dx} &= 1 \\ \frac{dy}{dx} &= \frac{1}{2y} = \frac{1}{2\left(-\sqrt{x}\right)} = -\frac{1}{2\sqrt{x}}. \end{aligned}$$

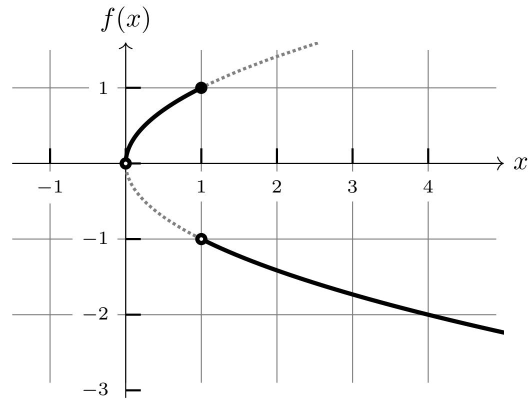
□

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

$$\begin{aligned}
qy^{q-1}\frac{dy}{dx} &= px^{p-1} \\
\frac{dy}{dx} &= \frac{px^{p-1}}{qy^{q-1}} \\
&= \frac{p}{q}\cdot\frac{x^{p-1}}{(x^r)^{q-1}} && \text{since } y = x^r \\
&= r\cdot x^{(p-1)-r(q-1)} \\
&= r\cdot x^{p-1-(p/q)(q-1)} \\
&= r\cdot x^{p-1-p+p/q} \\
&= r\cdot x^{-1+p/q} \\
&= r\cdot x^{r-1}.
\end{aligned}$$

□

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 Friday, 30 November)

Section 2.6, Problems 7–8 and 11–12. Problems 8 and 12 will be graded carefully.

Section 2.7, Problems 1–2. 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.)

MATH 131, LECTURE 24

CHARLES STAATS

1. Implicit Differentiation Note

One thing that was perhaps not clear from my previous discussion of implicit differentiation is the following:

> In implicit differentiation, you should often differentiate before doing anything else. Don't start solving for anything (or doing any algebra) until you have differentiated.

Example 1. Compute dy/dx in terms of x and y if $xy^2 + y^3 - y = 17\pi$.

Solution.

$$\begin{aligned}
xy^2 + y^3 - y &= 17\pi \\
\frac{d}{dx}\left(xy^2 + y^3 - y\right) &= \frac{d}{dx}(17\pi) && \text{differentiate both sides} \\
\left[1y^2 + x\frac{d}{dx}\left(y^2\right)\right] + \frac{d}{dx}\left(y^3\right) - \frac{d}{dx}(y) &= 0 && \text{product rule} \\
y^2 + x \cdot 2y\frac{dy}{dx} + 3y^2\frac{dy}{dx} - \frac{dy}{dx} &= 0 && \text{chain rule} \\
y^2 + \left(2xy + 3y^2 - 1\right)\frac{dy}{dx} &= 0 \\
\left(2xy + 3y^2 - 1\right)\frac{dy}{dx} &= -y^2 \\
\frac{dy}{dx} &= \frac{-y^2}{2xy + 3y^2 - 1}
\end{aligned}$$

□

2. Related Rates

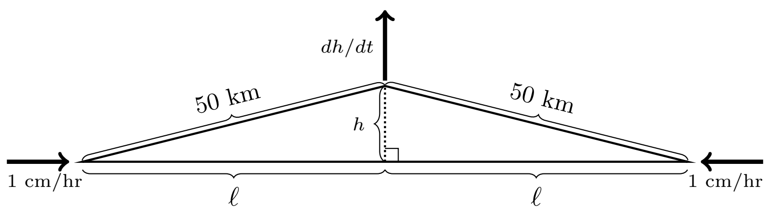
As far as I can tell, "related rates" are the textbook's first excuse to really start in on so-called "word problems." Up to now, the course has been mostly theoretical; the only real "applications" have been to studying the graphs of functions. However, calculus was invented for real-world problems. 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.

Date: Friday, 30 November 2012.

Example 2. 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

$$\begin{aligned} \left.\frac{dh}{dt}\right|_{t=1} &= \frac{1}{10^5 h} \\ &= \frac{1}{10^5 \cdot 10^{-2}\sqrt{1 - 10^{-6}}} \\ &= \frac{10^{-5+2}}{\sqrt{1 - 10^{-6}}} \\ &= \frac{.001}{\sqrt{1 - 10^{-6}}}. \end{aligned}$$

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, 3 December 2012)

Section 2.6, Problems 17 and 18. Problem 18 will be graded carefully.

Section 2.7, Problems 5–6 and 22. Problems 6 and 22 will be graded carefully.

Section 2.8, Problem 1.

Assignment 23 (due Wednesday, 5 December 2012)[1]

Section 2.7, Problem 9. This will be graded carefully.

Section 2.8, Problems 2, 3, and 6. Problems 2 and 6 will be graded carefully.

Section 3.1, Problems 1 and 5–6. On 5 and 6, include graphs of the functions on the interval. Do NOT graph the function outside the interval.

You will probably want to read Section 3.1 before (or while) attempting these three problems.

Differentiate the function f defined by

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

This will be graded carefully.

[1]This is the final homework assignment for this quarter.

MATH 131, LECTURE 25

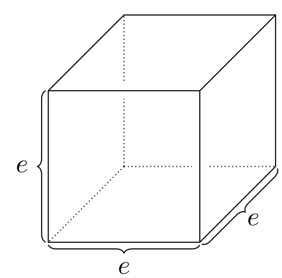
CHARLES STAATS

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.



Let's start with what is true "abstractly," i.e., for all time. The two quantities V (volume) and e (edge length) are related by the equation

$$V = e^3.$$

Differentiating implicitly, we see that

$$\frac{dV}{dt} = 3e^2 \frac{de}{dt}.$$

We also are given that, for all time, $de/dt = 3$. Substituting this in the equation above gives

$$\begin{aligned} &= 3e^2 \cdot 3 \\ &= 9e^2. \end{aligned}$$

(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

$$\frac{dV}{dt} = 9e^2 = 9(12)^2 = 9 \cdot 144 = 1296.$$

The volume is increasing at a rate of 1296 cubic inches per second. □

Date: Monday, 3 December 2012.

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).

$\square$

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

$$m = n \cdot p;$$

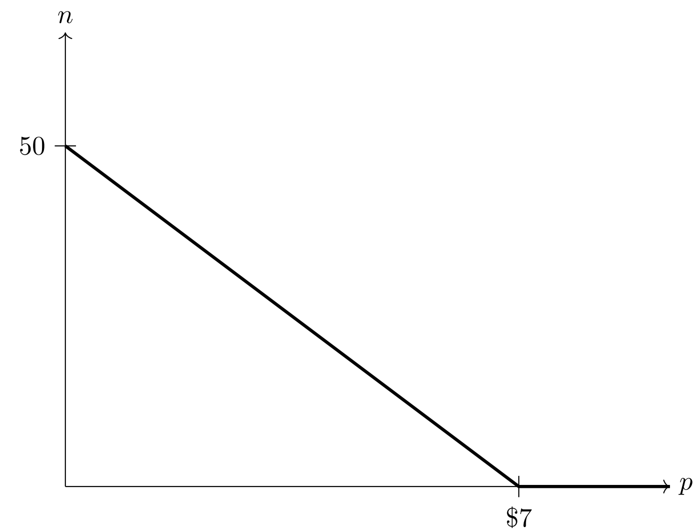
in words, the amount of money she makes is the number of cups she sells times the price per cup.[3] Moreover, we are assuming that the number of cups she sells is determined by the price she sets. In other words, n is a function of p. Consequently, m is also a function of p.

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.

[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.



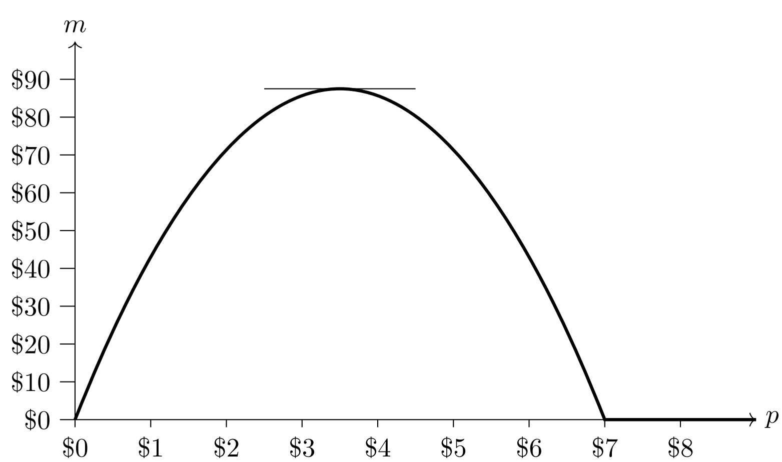
The corresponding function is

$$n = \begin{cases} 50 - \frac{50}{7}p & \text{if } 0 \le p \le 7, \\ 0 & \text{if } p > 7 \end{cases}$$

$$\begin{aligned} m = np &= \begin{cases} \left(50 - \frac{50}{7}p\right)p & \text{if } 0 \le p \le 7, \\ 0 \cdot p & \text{if } p > 7 \end{cases} \\ &= \begin{cases} 50p - \frac{50}{7}p^2 & \text{if } 0 \le p \le 7, \\ 0 & \text{if } p > 7. \end{cases} \end{aligned}$$

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 (due Wednesday, 5 December 2012)[4]

Section 2.7, Problem 9. This will be graded carefully.

Section 2.8, Problems 2, 3, and 6. Problems 2 and 6 will be graded carefully.

Section 3.1, Problems 1 and 5–6. On 5 and 6, include graphs of the functions on the interval. Do NOT graph the function outside the interval.

You will probably want to read Section 3.1 before (or while) attempting these three problems.

Differentiate the function f defined by

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

This will be graded carefully.

[4]This is the final homework assignment for this quarter.

MATH 131, LECTURE 26 (FINAL LECTURE)

CHARLES STAATS

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.

Annie and Ryan will not be holding tutorials tomorrow. Alex, however, will—in Ryan's usual tutorial room (HGS 361); everyone is welcome to attend, no matter whose tutorial you are normally assigned to.

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

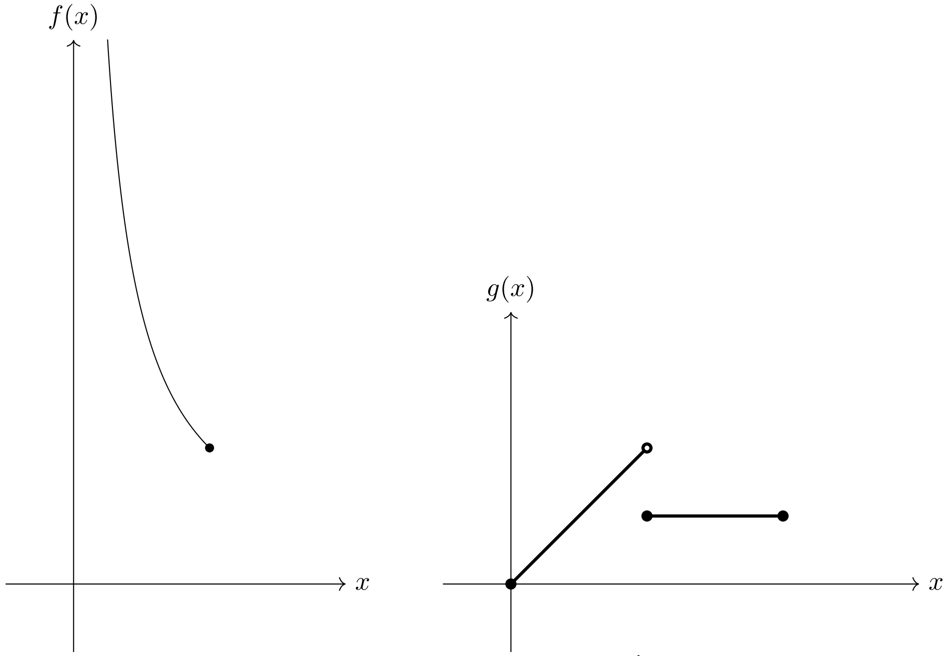
- What price should we sell lemonade at in order to make the most profit? (Profit is a function of price; we want to select price to maximize it.)
- What shape should a rectangle be to fence in the largest possible area with a fixed amount of fence? (The area of the rectangle is a function of its length; 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 an example problem.

Date: Wednesday, 5 December 2012.



$$f(x) = \frac{1}{x} \quad \text{if } 0 < x \leq 1$$

$$g(x) = \begin{cases} x & \text{if } 0 \leq x < 1, \\ \frac{1}{2} & \text{if } 1 \leq x \leq 2. \end{cases}$$

(A) This function has no maximum because it attains arbitrarily large values.

(B) 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 $g(x_0) = 1$.

FIGURE 1. Functions without maxima

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 cases in Figure 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,

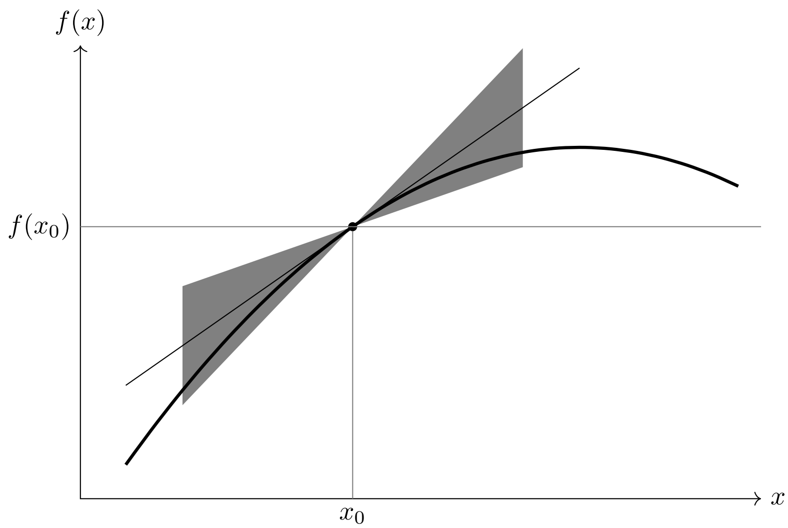
Definition. Let f be a function defined on an interval $[a, b]$ and x_0 a point in its domain. We say that x_0 is a *critical point* of f if any of the following holds:

- 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

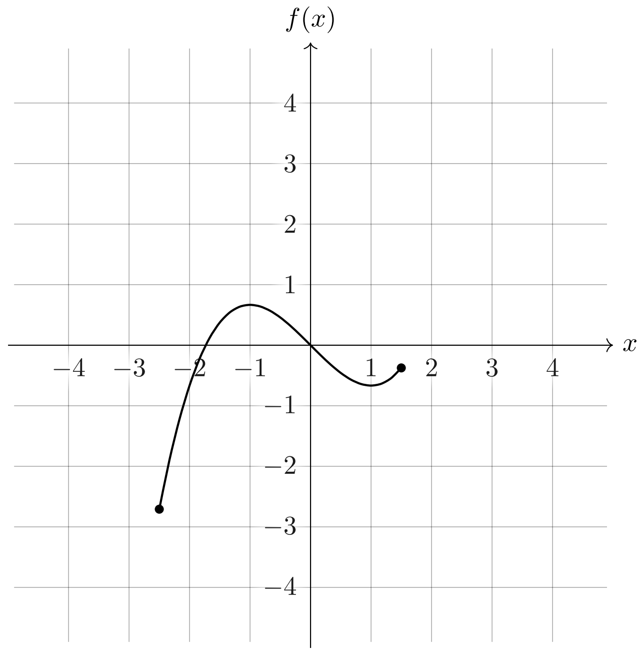


FIGURE 2. The function from Exercise 1

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 $[a, b]$.

Exercise 1. *Find the critical points, minimum, and maximum for the function f given by*

$$f(x) = \frac{1}{3}x^3 - x$$

on the closed interval $[-2.5, 1.5]$.

Solution. The graph of f is shown in Figure 2.

The endpoints -2.5 and 1.5 are critical points. The derivative

$$f'(x) = x^2 - 1$$

exists on the entire interval $(-2.5, 1.5)$, so there are no points at which f is not differentiable. To find the remaining critical points, we solve for zeros of f'.

$$\begin{aligned} & f'(x) = 0 \\ \iff \quad & x^2 - 1 = 0 \\ \iff \quad & (x-1)(x+1) = 0 \\ \iff \quad & x \in \{-1, 1\}. \end{aligned}$$

Therefore, the critical points of f are precisely -2.5, -1, 1, and 1.5. By the theorems above, we know that f has a maximum and minimum (since it is continuous on the closed interval $[-2.5, 1.5]$) and that the only possible values for these are

$$\begin{aligned} f(-2.5) &= -\tfrac{65}{24} = -\left(2 + \tfrac{17}{24}\right), \\ f(-1) &= \tfrac{2}{3}, \\ f(1) &= -\tfrac{2}{3}, \\ f(1.5) &= -\tfrac{3}{8}. \end{aligned}$$

Thus, the maximum is $\frac{2}{3}$ and the minimum is $-\frac{65}{24}$. $\square$