Welcome to 18.01

Welcome to 18.01. Today we start "Unit One"; the topic of the unit is differentiation. We'll start by reviewing what's in store in the next couple of weeks.

The topic of this lecture is "what is a derivative?" We're going to look at this question from several different points of view, and the first one is the geometric interpretation. We'll also discuss a physical interpretation.

Later we'll learn what makes calculus so fundamental in science and engineering. Derivatives are important in all measurements – in science, in engineering, in economics, in political science, in polling, in lots of commercial applications, in just about everything.

In this unit we'll also learn how to differentiate any function you know. That's a tall order, but by the end of the unit you will know how to take derivatives of functions like $f(x) = e^{x \cdot \arctan(x)}$.

Let's begin.

Geometric Interpretation of Differentiation

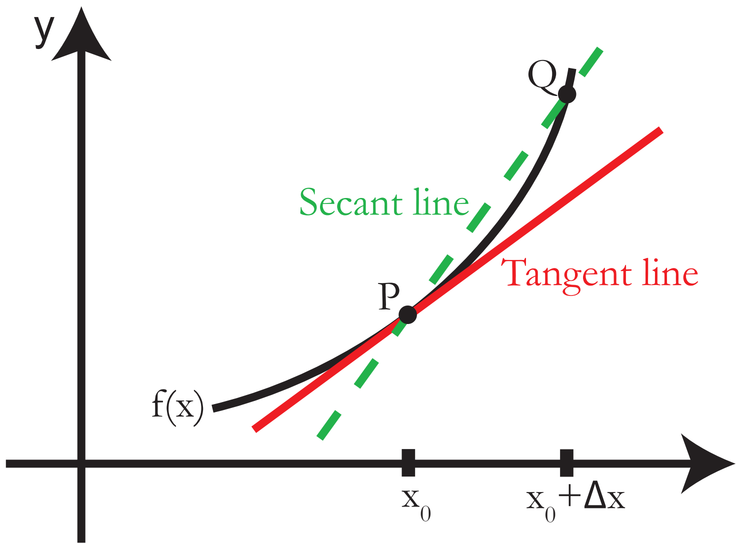


Figure 1: A graph with secant and tangent lines.

The derivative of $f(x)$ at $x = x_0$ is the slope of the tangent line to the graph of $f(x)$ at the point $(x_0, f(x_0))$. But what is a tangent line?

- It is NOT just a line that meets the graph at one point.
- It is the *limit* of the secant lines joining points $P = (x_0, f(x_0))$ and Q on the graph of $f(x)$ as Q approaches P.

The tangent line touches the graph at $(x_0, f(x_0))$; the slope of the tangent line matches the direction of the graph at that point. The tangent line is the straight line that best approximates the graph at that point.

Given a graph of our function, it's not hard for us to draw the tangent line to the graph. However, we'll want to do computations involving the tangent line and so will need a computational method of finding the tangent line.

How do we compute the equation of the line tangent to the graph of the function $f(x)$ at a point $P = (x_0, y_0)$? We know that the equation of the straight line with slope m through the point (x_0, y_0) is $y - y_0 = m(x - x_0)$, so in the abstract we know the equation of the tangent line.

To get a specific equation for the line, we'll need to know the coordinates x_0 and y_0 of the point P. If we know x_0 we can find $y_0 = f(x_0)$ by substituting the value x_0 in to the expression for $f(x)$. The second thing we need to know is the slope, $m = f'(x_0)$, which we call *the derivative of* f.

Definition: The derivative $f'(x_0)$ of f at x_0 is the slope of the tangent line to $y = f(x)$ at the point $P = (x_0, f(x_0)$.

Geometric definition of the derivative:

We're still trying to find a computational method of finding the equation of the tangent line – how do we compute the value of m?

In general, how do we know which lines are tangent lines and which lines are not?

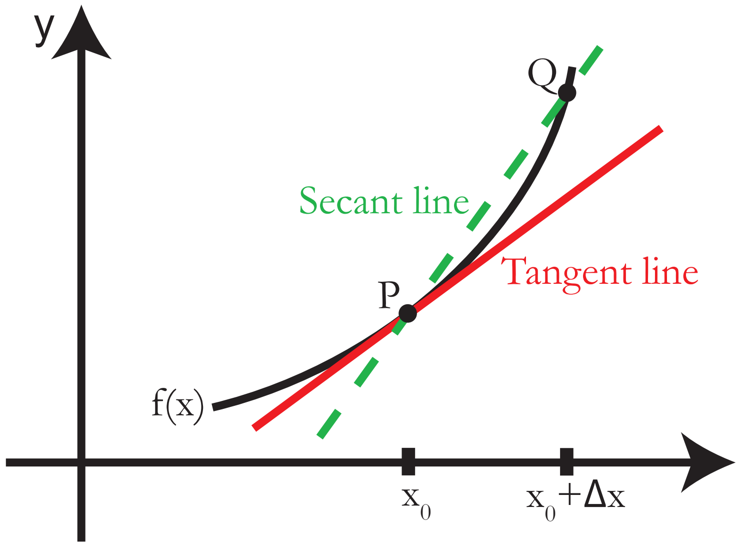


Figure 1: A graph with secant and tangent lines

A *secant* line is a line that joins two points on a curve. If the two points are close enough together, the slope of the secant line is close to the slope of the curve. We want to find the slope of the tangent line m — which equals the slope of the curve — and we use the slopes of secant lines to do this.

Suppose PQ is a secant line of the graph of $f(x)$. We can find the slope of the graph at P by calculating the slope of PQ as Q moves closer and closer to P (and the slope of PQ gets closer and closer to m).

The tangent line equals the limit of secant lines PQ as $Q \to P$; here P is fixed and Q varies.

Slope as Ratio

While we're still thinking geometrically, we can now use symbols and formulas in our computation.

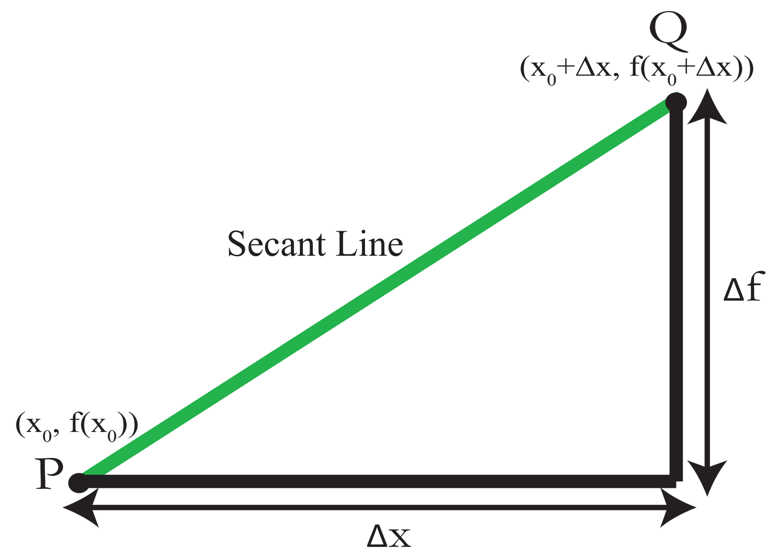


Figure 1: Geometric definition of the derivative

We start with a point $P = (x_0, f(x_0))$. We move over a tiny horizontal distance Δx (pronounced "delta x" and also called "the change in x") and find point $Q = (x_0 + \Delta x, f(x_0 + \Delta x))$. These two points lie on a secant line of the graph of $f(x)$; we will compute the slope of this line. The vertical difference between P and Q is $\Delta f = f(x_0 + \Delta x) - f(x_0)$.

The slope of the secant PQ is rise divided by run, or the ratio $\frac{\Delta f}{\Delta x}$. We've said that the tangent line is the limit of the secant lines. It is also true that the slope of the tangent line is the limit of the slopes of the secant lines. In other words,

$$m = \lim_{Q->P} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}.$$

Main Formula

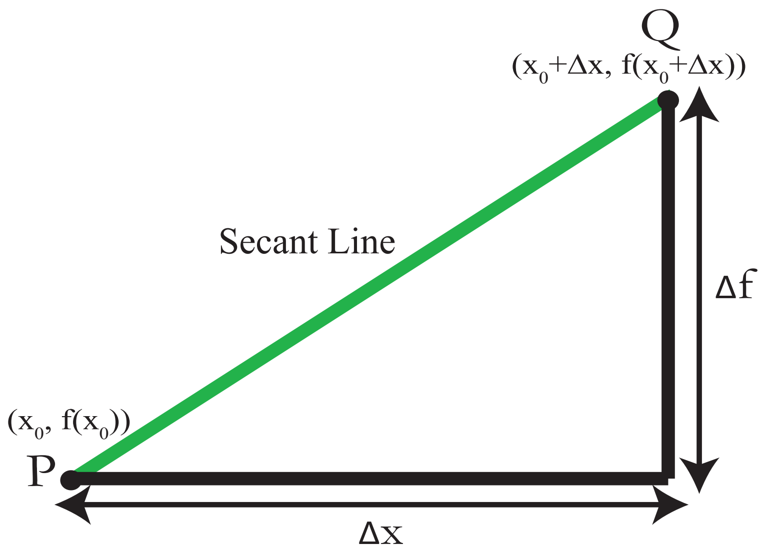


Figure 1: Geometric definition of the derivative

We started with a point P on the graph of $y = f(x)$ which had coordinates $(x_0, f(x_0))$. We then found a point Q on the the graph which was Δx units to the right of P. The coordinates of Q must be $(x_0 + \Delta x, f(x_0 + \Delta x))$. We can now write the following formula for the derivative:

$$m = \underbrace{f'(x_0)}_{\text{derivative of } f \text{ at } x_0} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \underbrace{\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}}_{\text{difference quotient}}$$

This is by far the most important formula in Lecture 1; it is the formula that we use to compute the derivative $f'(x_0)$, which equals the slope of the tangent line to the graph at P. A machine could use this formula together with the coordinates $(x_0, f(x_0))$ of the point P to draw the tangent line to the graph of $y = f(x)$ at the point P.

Example 1. $f(x) = \dfrac{1}{x}$

We'll find the derivative of the function $f(x) = \frac{1}{x}$. To do this we will use the formula:

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

Graphically, we will be finding the slope of the tangent line at at an arbitrary point $(x_0, \frac{1}{x_0})$ on the graph of $y = \frac{1}{x}$. (The graph of $y = \frac{1}{x}$ is a hyperbola in the same way that the graph of $y = x^2$ is a parabola.)

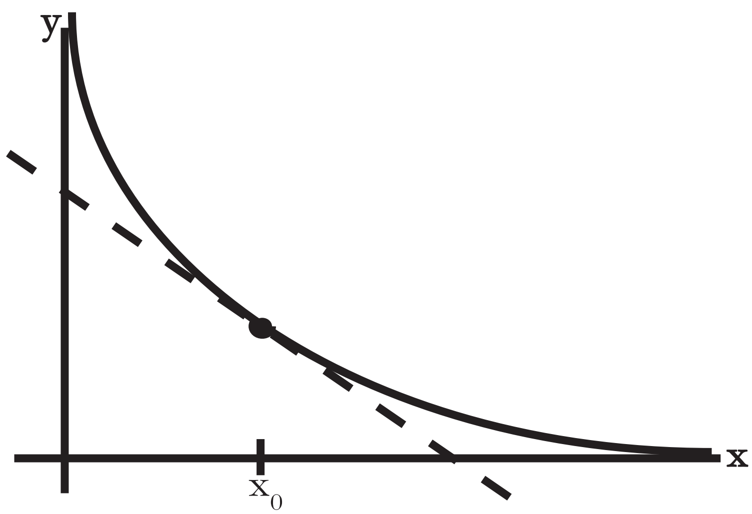


Figure 1: Graph of $\frac{1}{x}$

We start by computing the slope of the secant line:

$$\begin{aligned}
\frac{\Delta f}{\Delta x} &= \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \\
&= \frac{\frac{1}{x_0 + \Delta x} - \frac{1}{x_0}}{\Delta x} \\
&= \frac{(x_0)(x_0 + \Delta x)}{(x_0)(x_0 + \Delta x)} \frac{\frac{1}{x_0 + \Delta x} - \frac{1}{x_0}}{\Delta x} \\
&= \frac{\frac{(x_0)(x_0 + \Delta x)}{x_0 + \Delta x} - \frac{(x_0)(x_0 + \Delta x)}{x_0}}{(x_0)(x_0 + \Delta x)\Delta x} \\
&= \frac{1}{\Delta x} \frac{x_0 - (x_0 + \Delta x)}{(x_0)(x_0 + \Delta x)}
\end{aligned}$$

$$= \frac{1}{\Delta x}\frac{-\Delta x}{(x_0)(x_0+\Delta x)}$$
$$= \frac{-1}{(x_0)(x_0+\Delta x)}.$$

Next, we see what happens to the slopes of the secant lines as Δx tends to zero:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{-1}{(x_0)(x_0+\Delta x)} = \frac{-1}{x_0^2}$$

One thing to keep in mind when working with derivatives: it may be tempting to plug in $\Delta x = 0$ right away. If you do this, however, you will always end up with $\dfrac{\Delta f}{\Delta x} = \dfrac{0}{0}$. You will always need to do some cancellation to get at the answer.

We've computed that $f'(x) = \frac{-1}{x_0^2}$. Is this correct? How might we check our work? First of all, $f'(x_0)$ is negative — as is the slope of the tangent line on the graph of $y = \frac{1}{x}$. Secondly, as $x_0 \to \infty$ (i.e. as x_0 grows larger and larger), the tangent line is less and less steep. So $\frac{1}{x_0^2}$ should get closer to 0 as x_0 increases, which it does.

Question: Explain why $\displaystyle\lim_{\Delta x \to 0} \frac{-1}{(x_0)(x_0+\Delta x)} = \frac{-1}{x_0^2}$ again?

Answer: The point x_0 could be any point; let's suppose that $x_0 = 3$ so that we can look at this limit in a specific case.

We want to know the value of $\frac{-1}{(3)(3+\Delta x)}$ as Δx tends toward zero. As Δx gets smaller and smaller $3 + \Delta x$ gets closer and closer to 3, and so $\frac{-1}{(3)(3+\Delta x)}$ gets closer and closer to $\frac{-1}{(3)(3)} = \frac{-1}{9}$.

Question: Why is it that $\dfrac{\frac{1}{x_0+\Delta x} - \frac{1}{x_0}}{\Delta x} = \dfrac{1}{\Delta x}\dfrac{x_0 - (x_0+\Delta x)}{(x_0)(x_0+\Delta x)}$?

Answer: There are two steps in this simplification. We factored out the Δx that was in the denominator to become the $\frac{1}{\Delta x}$ "out front". At the same time, we rewrote the difference of two fractions $\frac{1}{x_0+\Delta x} - \frac{1}{x_0}$ using a common denominator.

This common denominator was $(x_0)(x_0 + \Delta x)$, which is just the product of the denominators in $\frac{1}{x_0+\Delta x} - \frac{1}{x_0}$. To get the common denominator, we multiply the first fraction by $\frac{(x_0)}{(x_0)} = 1$ and the second by $\frac{(x_0+\Delta x)}{(x_0+\Delta x)}$. (Multiplying by 1 won't change its value, but can change the algebraic expression we use to describe that value.) The denominators cancel, as intended, and we're left with $\frac{x_0-(x_0+\Delta x)}{(x_0)(x_0+\Delta x)}$.

A "Harder" Problem

People say that calculus is hard, but the example we just saw — computing the derivative of $f(x) = \frac{1}{x}$ — was not very difficult. What makes calculus seem hard is the context calculus problems appear in. For example, the problem we are about to solve combines algebra, geometry and problem solving with calculus. Because we use calculus to solve it, it is "a calculus problem". And although it is a harder problem, it's not the calculus that makes it hard.

So far all we've talked about is geometry, so our example problem must be geometric.

Problem: Find the area of the triangle formed by the x and y-axes and the tangent to the graph of $y = \frac{1}{x}$.

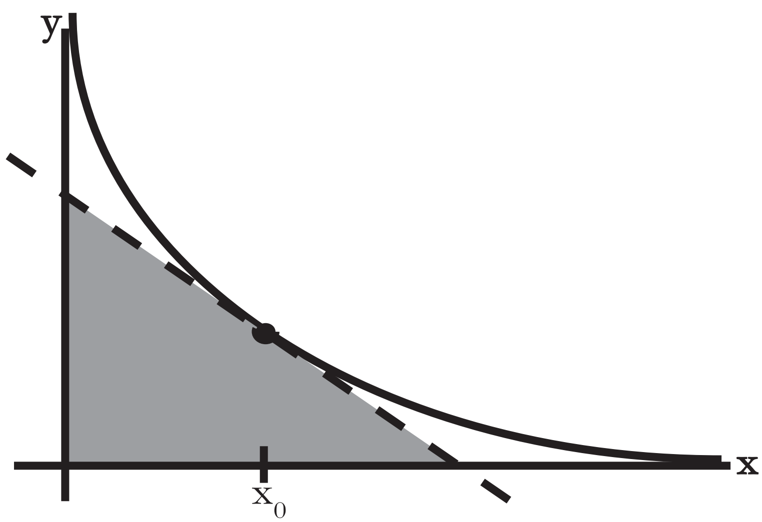


Figure 1: Triangle formed by axes and tangent line

We start by drawing a picture. As we draw the picture, we realize we've assumed that the triangle lies in the first quadrant. The solution we find might or might not depend on this assumption — we'll have to do some work later if we want to be sure it's true when x and $\frac{1}{x}$ are negative.

Because the problem refers to a tangent line it is a calculus problem, but as you'll see, the calculus is the easy part.

The next step in our solution is labeling the picture. We label the graph $y = \frac{1}{x}$. We'd like to put some labels on the triangle, like the lengths of its sides, but we don't know how to find those numbers. What we do know is that the hypotenuse of the triangle is tangent to the graph, and we can label the point of tangency $(x_0, \frac{1}{x_0})$ (which lets us label the points $(x_0, 0)$ and $(0, y_0)$ on the axes).

The area of a triangle is $\frac{1}{2}b \cdot h$, so we'd like to find the lengths of the base and height of this triangle. If we can find the x-coordinate of the point where the tangent line intersects the x-axis, we'll know the length of the base of the triangle. Similarly, finding the y-intercept of the tangent line will give us the triangle's height.

In order to find the x- and y-intercepts of the tangent line, we first must find the equation of the line. We start by writing down the "point-slope" form of the equation for a line:

$$y - y_0 = m(x - x_0).$$

The part of this problem that requires calculus is finding the slope m of the tangent line. Luckily, we already did this calculation in the previous example: $m = -\frac{1}{x_0^2}$.

$$y - y_0 = -\frac{1}{x_0^2}(x - x_0)$$

We've finished all the calculus in the problem, but we still need to do some work to find the area of the triangle. The point (x_0, y_0) is the point where the hypotenuse is tangent to the graph of $y = \frac{1}{x}$. We leave x_0 as is, and replace y_0 by $\frac{1}{x_0}$.

$$y - \frac{1}{x_0} = -\frac{1}{x_0^2}(x - x_0).$$

First we calculate the x-intercept of the tangent line, which will give us the length of the base of the triangle. The line crosses the x-axis when $y = 0$. Setting $y = 0$ in the equation for the tangent line, we get:

$$\begin{aligned}
0 - \frac{1}{x_0} &= \frac{-1}{x_0^2}(x - x_0) \\
\frac{-1}{x_0} &= \frac{-1}{x_0^2}x + \frac{1}{x_0} \\
\frac{1}{x_0^2}x &= \frac{2}{x_0} \\
x &= x_0^2\left(\frac{2}{x_0}\right) = 2x_0
\end{aligned}$$

So, the x-intercept of this tangent line is at $x = 2x_0$.

Next we could find the y-intercept using a very similar calculation — replacing x by 0 and solving for y (because the y-intercept is the point on the line with x coordinate 0). If we don't want to do the calculation all over again, or if we like using clever tricks to make problems easier, we can use the symmetry of the graph to take a short cut.

Since $y = \dfrac{1}{x}$ and $x = \dfrac{1}{y}$ are identical equations, the graph of $y = \frac{1}{x}$ is symmetric when x and y are exchanged. By symmetry, then, we could swap the

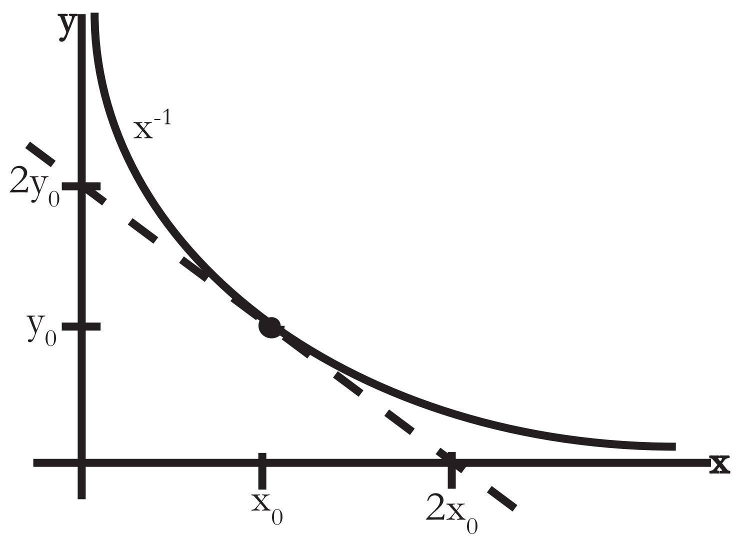


Figure 2: Triangle formed by axes and tangent line, labeled

x's in the calculation above with the y's to conclude that the y-intercept is at $y = 2y_0 = \frac{2}{x_0}$.

Finally,

$$\text{Area} = \frac{1}{2}\text{b} \cdot \text{h} = \frac{1}{2}(2\text{x}_0)(2\text{y}_0) = 2\text{x}_0\text{y}_0 = 2\text{x}_0(\frac{1}{\text{x}_0}) = 2 \text{ (see Fig. 2)}$$

Curiously, the area of the triangle is *always* 2, no matter where on the graph we draw the tangent line!

Remark: We call it "one variable calculus", but we just used *four* variables: x, y, x_0, and y_0. We could have had more! This makes things complicated, and it's something that you'll have to get used to.

Another complicated thing that we do is reuse variables. In this problem, there are (at least) three different possible interpretations of the variable y. When we said $y = \frac{1}{x}$, we were thinking of y as the vertical position of a point on the hyperbola. When we said $y - y_0 = \frac{-1}{x_0^2}(x - x_0)$ we were thinking of y as the vertical position of a point on the tangent line. And when we said $y = 0$ we were talking about the vertical position of all the points on the x-axis. Once you've practiced calculus for a while you will know from context which meaning y has, just as you can tell in conversation whether a person is saying "sea" or "see". Until you reach that point, be sure you understand the meaning of an equation before using it in a calculation.

Notations

Calculus, rather like English or any other language, was developed by several people. As a result, just as there are many ways to express the same thing, there are many notations for the derivative.

Since $y = f(x)$, it's natural to write

$$\Delta y = \Delta f = f(x) - f(x_0) = f(x_0 + \Delta x) - f(x_0)$$

We say "Delta y" or "Delta f" or the "change in y".

If we divide both sides by $\Delta x = x - x_0$, we get two expressions for the difference quotient:

$$\frac{\Delta y}{\Delta x} = \frac{\Delta f}{\Delta x}$$

Taking the limit as $\Delta x \to 0$, we get

$$\begin{array}{ccl} \dfrac{\Delta y}{\Delta x} & \to & \dfrac{dy}{dx} \text{ (Leibniz' notation)} \\ \dfrac{\Delta f}{\Delta x} & \to & f'(x_0) \text{ (Newton's notation)} \end{array}$$

In Leibniz' notation we might also write $\frac{df}{dx}$, $\frac{d}{dx}f$ or $\frac{d}{dx}y$. Notice that Leibniz' notation doesn't specify where you're evaluating the derivative. In the example of $f(x) = \frac{1}{x}$ we were evaluating the derivative at $x = x_0$.

Other, equally valid notations for the derivative of a function f include f' and Df.

Example 2. $f(x) = x^n$ where $n = 1, 2, 3...$

In this example we answer the question "What is $\frac{d}{dx}x^n$?" Once we know the answer we can use it to, for example, find the derivative of $f(x) = x^4$ by replacing n by 4.

At this point in our studies, we only know one tool for finding derivatives – the difference quotient. So we plug $y = f(x)$ into the definition of the difference quotient:

$$\frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{(x_0 + \Delta x)^n - x_0^n}{\Delta x}$$

Because writing little zeros under all our x's is a nuisance and a waste of chalk (or of photons?), and because there's no other variable named x to get confused with, from here on we'll replace x_0 with x.

$$\frac{\Delta y}{\Delta x} = \frac{(x + \Delta x)^n - x^n}{\Delta x}$$

Remember that when we use the difference quotient, we're thinking of x as fixed and of Δx as getting closer to zero. We want to simplify this fraction so that we can plug in 0 for Δx without any danger of dividing by zero. To do this we must expand the expression $(x + \Delta x)^n$.

A famous formula called the binomial theorem tells us that:

$$(x + \Delta x)^n = (x + \Delta x)(x + \Delta x)...(x + \Delta x) \quad \text{n times}$$

We can rewrite this as

$$x^n + n(\Delta x)x^{n-1} + O\left((\Delta x)^2\right)$$

where $O(\Delta x)^2$ is shorthand for "all of the terms with $(\Delta x)^2$, $(\Delta x)^3$, and so on up to $(\Delta x)^n$."

One way to begin to understand this is to think about multiplying all the x's together from

$$(x + \Delta x)^n = (x + \Delta x)(x + \Delta x)...(x + \Delta x) \quad \text{n times.}$$

There are n of these x's, so multiplying them together gives you one term of x^n. What if you only multiply together $n - 1$ of the x's? Then you have one $(x+\Delta x)$ left that you haven't taken an x from, and you can multiply your x^{n-1} by Δx. (If you multiplied by x, you'd just have the x^n that you already got.) There were n different Δx's that you could have chosen to use, so you can get this result n different ways. That's where the $n(\Delta x)x^{n-1}$ comes from.

We could keep going, and figure out how many different ways there are to multiply $n - 2$ x's by two Δx's, and so on, but it turns out we don't need to. Every other way of multiplying together one thing from each $(x + \Delta x)$ gives you at least two Δx's, and $\Delta x \cdot \Delta x$ is going to be too small to matter to us as $\Delta x \to 0$.

Now that we have some idea of what $(x + \Delta x)^n$ is, let's go back to our difference quotient.

$$\frac{\Delta y}{\Delta x} = \frac{(x+\Delta x)^n - x^n}{\Delta x} = \frac{(x^n + n(\Delta x)(x^{n-1}) + O(\Delta x)^2) - x^n}{\Delta x} = nx^{n-1} + O(\Delta x)$$

As it turns out, we *can* simplify the quotient by canceling a Δx in all of the terms in the numerator. When we divide a term that contains Δx^2 by Δx, the Δx^2 becomes Δx and so our $O(\Delta x^2)$ becomes $O(\Delta x)$.

When we take the limit as x approaches 0 we get:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = nx^{n-1}$$

and therefore,

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

This result is sometimes called the "power rule". We will use it often to find derivatives of polynomials; for example,

$$\frac{d}{dx}(x^2 + 3x^{10}) = 2x + 30x^9$$

Introduction to Rates of Change

Last class, we defined the derivative as the slope of a tangent line. Today we'll see how to interpret the derivative as a rate of change, clarify the idea of a limit, and use this notion of limit to describe continuity – a property functions need to have in order for us to work with them.

Rates of Change

Last class we talked about the derivative as the slope of the tangent line to a graph. This class we'll continue our discussion of derivatives by explaining how a derivative can be a rate of change. This some of the most important information presented in this class.

Remember that when we talked about the slope of a graph $y = f(x)$ we started by talking about the change in y and the change in x. If changing x at a certain rate causes y to change, we're interested in the *relative* rate of change, $\frac{\Delta y}{\Delta x}$.

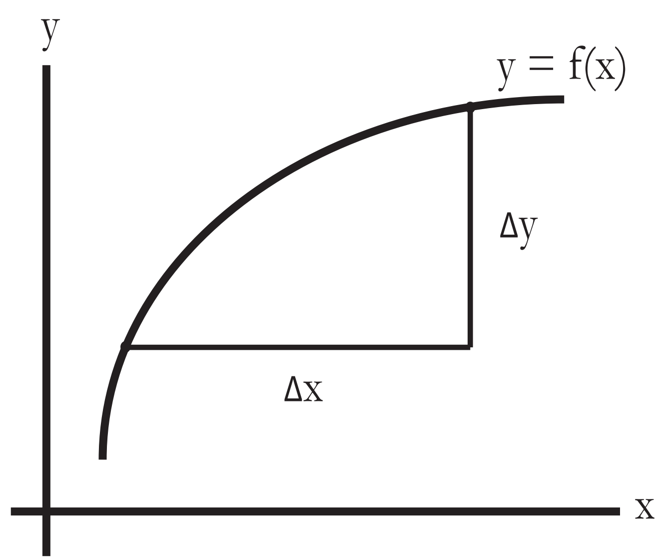


Figure 1: Graph of a generic function, with Δx and Δy marked on the graph

Another way to think about $\frac{\Delta y}{\Delta x}$ is as the average change in y over an interval of size Δx. This comes up frequently in physics, in which x is measuring time and $\frac{\Delta y}{\Delta x}$ is the average change in position over an interval of time – in other words, it's the rate at which something is moving. In this case, the limit

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

measures the instantaneous rate of change, or the speed.

Physical Interpretation of Derivatives

You can think of the derivative as representing a rate of change (speed is one example of this). This makes it very useful for solving physics problems.

Here's one example from physics: If q is an amount of electric charge, the derivative $\frac{dq}{dt}$ is the change in that charge over time, or the electric current.

A second, more tangible example is to let s stand for distance; then the rate of change $\frac{ds}{dt}$ is what we call speed. Let's investigate this second example in more detail to get a visceral sense of what instantaneous speed means.

On Halloween, MIT students have a tradition of dropping pumpkins from the roof of the building this lecture was given in. Let's say that the building is about 300 feet tall. We'll use a slightly smaller value of 80 meters for the height because it makes the problem easier to solve.

The equation of motion for objects near the earth's surface (which we will just accept for now) says that the height above the ground h of the pumpkin t seconds after it's dropped from the building is roughly:

$$h = 80 - 5t^2 \text{ meters}$$

Let's think about this. The instant the pumpkin is dropped, $t = 0$ and $h = 80$ meters. When $t = 4$ seconds, $h = 80 - 5(4^2) = 0$, and the pumpkin has reached the ground.

The average speed of the pumpkin over the time it's falling is

$$\frac{\Delta h}{\Delta t} = \frac{\text{distance traveled}}{\text{time elapsed}} = \frac{0 - 80 \text{ meters}}{4 - 0 \text{ seconds}} = -20 \text{ m/s}.$$

(The numerator is $0 - 80$ and not $80 - 0$ because we must subtract the initial position from the final position, not the other way around.)

The people watching the pumpkin drop probably don't care about the *average* speed. They want to know how fast the pumpkin is going when it slams into the ground. That's known as the instantaneous speed, and is the derivative $h'(t) = \frac{d}{dt}h$. To find the instantaneous velocity at $t = 5$, we evaluate $\frac{d}{dt}h$.

$$\frac{d}{dt}h = 0 - 10t = -10t$$

If you've had calculus before, you're probably able to find the derivative of the polynomial $80 - 5t^2$ on your own. If not, you'll have to take a few things on faith here. First, the derivative of $80 - 5t^2$ is just the derivative of 80 minus the derivative of $5t^2$. Next, the derivative of 80 is the slope of the graph of $y = 80$ when $x = 0$; that graph is a horizontal line! And finally, since we know from last class that the derivative of t^2 is $2t^{2-1} = 2t$ it should not surprise you that the derivative of $5t^2$ is $10t$.

We know that the pumpkin hits the pavement 4 seconds after it's dropped, at time $t = 4$, so the pumpkin's speed is:

$$h'(4) = (-10)(4) = -40 \text{ m/s} \quad \text{(about 90 mph or 145 kph)}.$$

The value of $\frac{d}{dt}h$ is negative because the pumpkin's height is decreasing; it is moving downward.

In actuality, the building is a little taller than 80 meters and there is air resistance. You may do a much more thorough study on your own if you wish.

Physical Interpretation of Derivatives, Continued

Calculus is a good tool for studying how things change over time; we can also use it to investigate change with respect to variables other than time. Let's look at a couple of examples that don't involve time as a variable.

If we let T denote temperature and x measure a distance or position, then $\frac{dT}{dx}$ is the rate the temperature changes as position changes. This quantity is called the *temperature gradient*. Temperature gradients are important in weather forecasting because it's that temperature difference that causes air flows and weather changes.

$T =$ temperature $\dfrac{dT}{dx} =$ temperature gradient

Another application of calculus is to "sensitivity of measurements". In Problem Set 1 there's a question about GPS and a satellite which explores sensitivity of measurements.

In the problem set, you assume that the earth is flat and that you have a satellite above a known location. A traveler's GPS device measures the distance h between the traveler and the satellite and uses that information to compute the horizontal distance L between the traveler and the point directly under the satellite. (See Fig. 1 and Fig. 2)

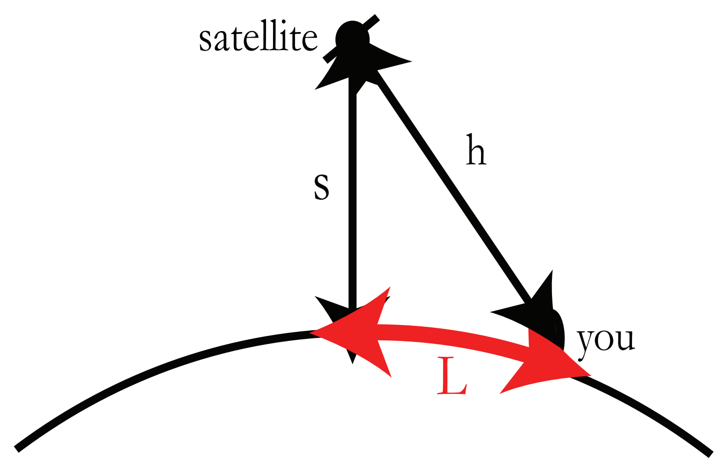


Figure 1: The Global Positioning System Problem (GPS)

In other words, the GPS computes L as a function of h. But there is usually some error Δh in your measurement of h; given that, how accurately can we measure L? The error ΔL is estimated by looking at $\dfrac{\Delta L}{\Delta h} \approx \dfrac{dL}{dh}$.

Why is this important? For one thing, it's used all the time to land airplanes.

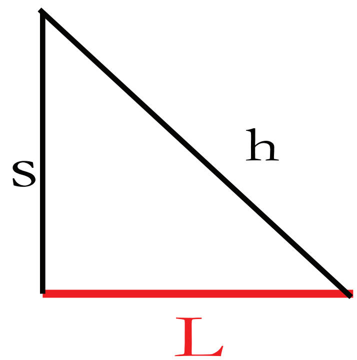


Figure 2: On problem set 1, you will look at this simplified “flat earth” model

This concludes our introduction to differentiation, although there will be plenty of opportunities throughout the course for you to improve your understanding.

Limits

Last class we talked about a series of secant lines approaching the "limit" of a tangent line, and about how as Δx approaches zero, $\frac{\Delta y}{\Delta x}$ approaches the "limit" $y' = \frac{dy}{dx}$. Now we want to talk about limits more carefully; this will include some of our first steps towards our goal of being able to differentiate every function you know.

Some limits are easy to compute:

$$\lim_{x \to 3} \frac{x^2 + x}{x + 1} = \frac{3^2 + 3}{3 + 1} = \frac{12}{4} = 3$$

With an easy limit, you can get a meaningful answer just by plugging in the limiting value. This is because when x is close to 3, the value of the function $f(x) = \frac{x^2+x}{x+1}$ is close to $f(3)$.

Some limits are not easy to compute. For example, the definition of the derivative:

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

is never an easy limit, because the denominator $\Delta x = 0$ is not allowed. (The limit $x \to x_0$ is computed under the implicit assumption that $x \neq x_0$.) We'll always need to cancel Δx before we can make sense out of the limit.

Other "hard" limits would be:

$$\lim_{x \to -1} \frac{x^2 + x}{x + 1} \quad \text{and} \quad \lim_{x \to \infty} \frac{x^2 + x}{x + 1}.$$

Any limit involving infinity or division by zero is going to be harder to compute; sometimes the answer will will be that there is no limit.

To complete our discussion of limits, we need just one more piece of notation — the concepts of left hand and right hand limits.

The limit

$$\lim_{x \to x_0^+} f(x)$$

is known as the *right-hand limit* and means that you should use values of x that are greater than x_0 (to the right of x_0 on the number line) to compute the limit. Shown below is the graph of the function:

$$f(x) = \begin{cases} x + 1 & x > 0 \\ -x & x \leq 0 \end{cases}$$

The right-hand limit $\lim_{x \to 0^+} f(x)$ equals 1.

The *left-hand limit*

$$\lim_{x \to x_0^-} f(x)$$

is found by looking at values of $f(x)$ when x is less than x_0 (to the left of x_0 on the number line). For this function, $\lim_{x \to 0^-} f(x) = 0$.

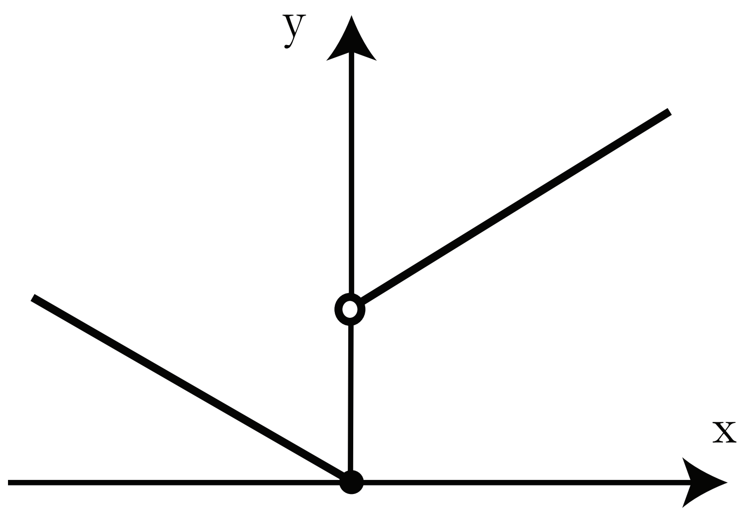


Figure 1: Graph of $f(x)$

The notions of left- and right- hand limits will make things much easier for us as we discuss continuity, next.

Let's talk more about the example graphed above. To calculate

$$\lim_{x \to x_0^+} f(x)$$

we use only values of x that are greater than 0. When $x > 0$, $f(x)$ is defined to equal $x + 1$. So we plugged $x = 0$ into the expression $x + 1$ to calculate the right-hand limit.

When calculating

$$\lim_{x \to x_0^-} f(x),$$

we have $x < 0$. Here $f(x)$ is defined to equal $-x$; when we plug $x = 0$ into this expression we get $\lim_{x \to x_0^-} f(x) = 0$.

Notice that it doesn't matter that $f(0) = 0$. Our calculations would have been exactly the same if $f(0)$ were 1 or even if $f(0) = 2$.

Continuity

Continuous Functions

Definition: A function f is *continuous* at x_0 if $\lim_{x\to x_0} f(x) = f(x_0)$.

What is this definition saying? A function that's continuous at x_0 has the following properties:

- $\lim_{x\to x_0^+} f(x) = \lim_{x\to x_0^-} f(x)$; in particular, both of these one sided limits exist.
- $f(x_0)$ is defined.
- $\lim_{x\to x_0^+} f(x) = \lim_{x\to x_0^-} f(x) = f(x_0)$.

This may look obvious, but remember that when you are calculating $\lim\limits_{x\to x_0} f(x)$ you never allow x to equal x_0. The value $\lim\limits_{x\to x_0} f(x)$ is computed independently of, and in a different way than, the value of $f(x_0)$. If we aren't careful to make this distinction, this definition has no meaning.

The limits for which $\lim\limits_{x\to x_0} f(x) = f(x_0)$ are exactly the "easy limits" we discussed earlier. The "harder" limits only happen for functions that are not continuous.

Next we'll see a tour of different types of discontinuous functions. The question of whether something is continuous or not may seem fussy, but it is something people have worried about a lot. Bob Merton, who was a professor at MIT when he did his work for the Nobel Prize in Economics, was interested in whether stock prices of various kinds are continuous from the left (past) or right (future) in a certain model. That was a serious consideration when developing a model that hedge funds now use all the time.

Jump Discontinuity

A *jump* discontinuity occurs when the right-hand and left-hand limits exist but are not equal. We've already seen one example of a function with a jump discontinuity:

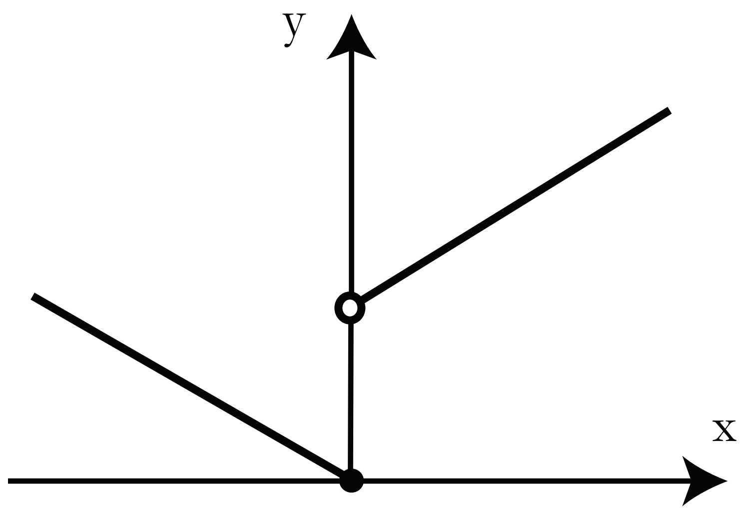


Figure 1: Graph of the discontinuous function listed below

$$f(x) = \begin{cases} x+1 & x>0 \\ -x & x \geq 0 \end{cases}$$

This *discontinuous* function is seen in Fig. 1. For $x > 0$,

$$\lim_{x \to 0} f(x) = 1$$

but $f(0) = 0$. (One can also say, f is continuous from the left at 0, but not the right.)

Here is another example in which $\lim\limits_{x \to x_0^+}$ exists, and $\lim\limits_{x \to x_0^-}$ also exists, but they are NOT equal.

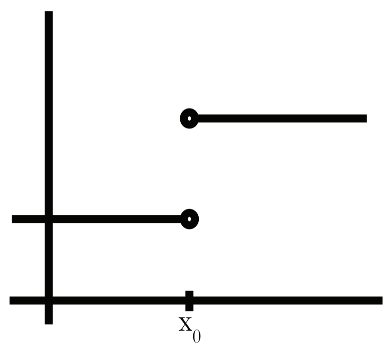


Figure 2: Another example of a jump discontinuity

Removable Discontinuities

At a *removable* discontinuity, the left-hand and right-hand limits are equal but either the function is not defined or not equal to these limits:

$$\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x) \neq f(x_0)$$

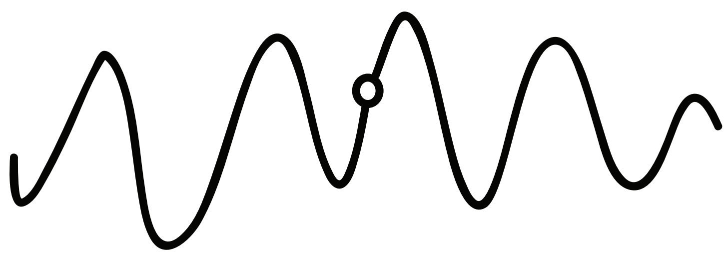


Figure 1: A removable discontinuity: the function is continuous everywhere except one point

For example, $g(x) = \dfrac{\sin(x)}{x}$ and $h(x) = \frac{1-\cos x}{x}$ are defined for $x \neq 0$, but both functions have removable discontinuities. This is not obvious at all, but we will learn later that:

$$\lim_{x \to 0} \frac{\sin x}{x} = 1 \text{ and } \lim_{x \to 0} \frac{1 - \cos x}{x} = 0.$$

So both of these functions have removable discontinuities at $x = 0$ despite the fact that the fractions defining them have a denominator of 0 when $x = 0$.

Infinite Discontinuities

In an *infinite* discontinuity, the left- and right-hand limits are infinite; they may be both positive, both negative, or one positive and one negative.

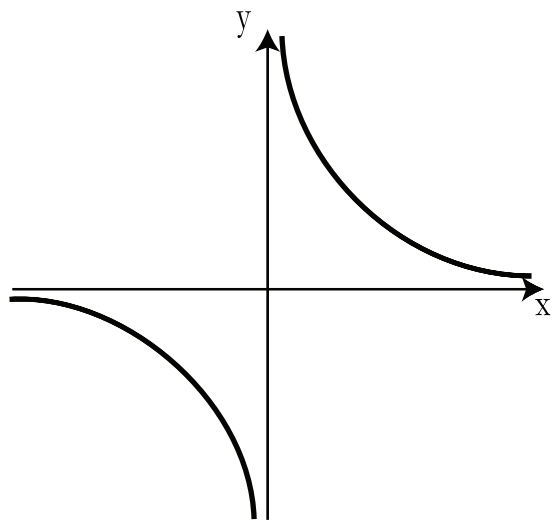


Figure 1: An example of an infinite discontinuity: $\dfrac{1}{x}$

From Figure 1, we see that $\lim\limits_{x\to 0^+} \dfrac{1}{x} = \infty$ and $\lim\limits_{x\to 0^-} \dfrac{1}{x} = -\infty$. Saying that a limit is ∞ is different from saying that the limit doesn't exist – the values of $\frac{1}{x}$ are changing in a very definite way as $x \to 0$ from either side. (Note that it's not true that $\lim\limits_{x\to 0} \dfrac{1}{x} = \infty$ because ∞ and $-\infty$ are different.)

There are two more things we can learn from this example. First, sketch the graph of $\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$; it also has an infinite discontinuity at $x = 0$. Notice that the derivative of the function $\frac{1}{x}$ is always negative. It may seem strange to you that the derivative is decreasing as x approaches 0 from the positive side while $\frac{1}{x}$ is increasing, but very often the graph of the derivative will look nothing like the graph of the original function.

What the graph of the derivative $-\frac{1}{x^2}$ is showing you is the slope of the graph of $\frac{1}{x}$. Where the graph of $\frac{1}{x}$ is not very steep, the graph of $-\frac{1}{x^2}$ lies close to the x-axis. Where the graph of $\frac{1}{x}$ is steep, the graph of $-\frac{1}{x^2}$ is far away from the x-axis. The value of $-\frac{1}{x^2}$ is always negative, and the graph of $\frac{1}{x}$ always slopes downward.

Finally, $\frac{1}{x}$ is an odd function and $-\frac{1}{x^2}$ is an even function. When you take the derivative of an odd function you always get an even function and vice-versa. If you can easily identify odd and even functions, this is a good way to check

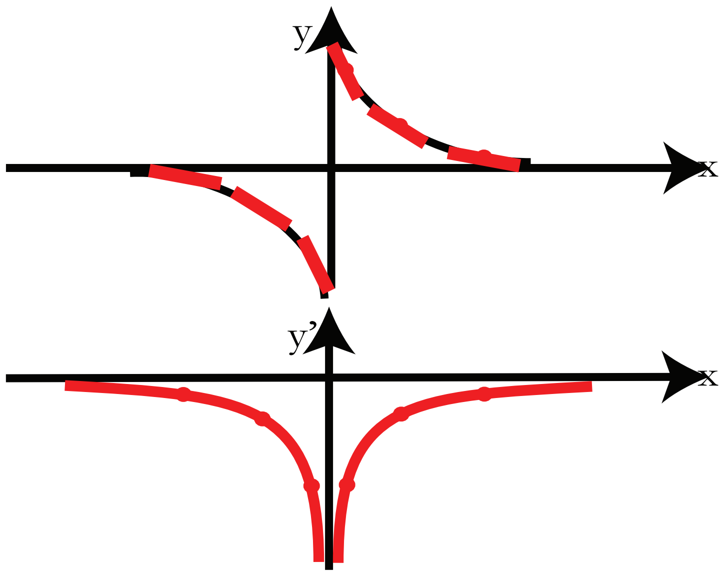


Figure 2: Top: graph of $f(x) = \dfrac{1}{x}$ and Bottom: graph of $f'(x) = -\dfrac{1}{x^2}$

your work.

Other (Ugly) Discontinuities

The limit $\lim_{x \to 0} \sin(1/x)$ *is* undefined as x goes to 0. The graph of $y = \sin(1/x)$ is similar to the one in Figure 1; the function $\sin(1/x)$ has no left or right limit as x goes to 0. Here, we say the limit does not exist.

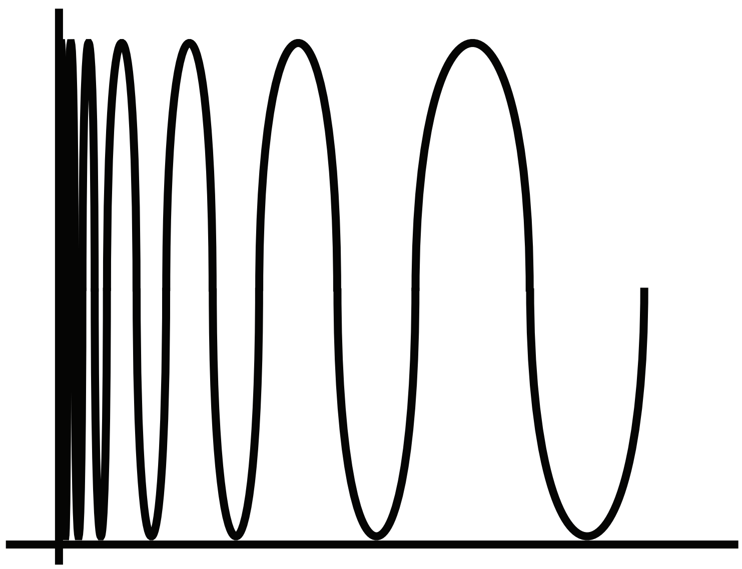


Figure 1: An example of an ugly discontinuity: a function that oscillates a lot as it approaches the origin

There are many discontinuities of this type — for example, things that oscillate as time goes to infinity — but we're not going to worry about them in this course.

Differentiable Implies Continuous

Theorem: If f is differentiable at x_0, then f is continuous at x_0.

We need to prove this theorem so that we can use it to find general formulas for products and quotients of functions.

We begin by writing down what we need to prove; we choose this carefully to make the rest of the proof easier. We want to show that:

$$\lim_{x \to x_0} f(x) - f(x_0) = 0.$$

This is the same as saying that the function is continuous, because to prove that a function was continuous we'd show that $\lim_{x \to x_0} f(x) = f(x_0)$.

We prove $\lim_{x \to x_0} f(x) - f(x_0) = 0$ by multiplying and dividing it by the same number – this won't change its value.

$$\begin{aligned} \lim_{x \to x_0} f(x) - f(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \\ &= f'(x) \cdot 0 \\ &= 0. \end{aligned}$$

(Notice that we used our assumption that f was differentiable when we wrote down $f'(x)$.)

But wait! When we multiplied and divided by $x - x_0$ weren't we multiplying and dividing by zero? We know from our algebra classes that this never works! It turns out that we're safe because we're using limits. Although x gets closer and closer to x_0, it never actually equals x_0, and so we never quite divide by 0. That's what limits are for; $x - x_0$ may be small, but it's always non-zero.

So this calculation is valid, it's true that $\lim_{x \to x_0} f(x) - f(x_0) = 0$, and it's true that differentiable functions are continuous.

Introduction to Differentiation

Working toward our goal of "differentiating everything", this lecture introduces some useful new formulas.

There are two basic types of derivative formulas:

1. Specific Examples: power rule

2. General Examples: (u +v)' = u'+v' and (cu)' = cu' (where c is a constant)

We need both kinds of formulas to take derivatives of polynomials, for example.

This lecture focuses on the basic trig functions, finding specific formulas for the derivative of the sine function and the cosine function.

Derivative of a Sum

One of our examples of a general derivative formula was:

$$(u+v)'(x) = u'(x) + v'(x).$$

(Remember that by $(u+v)(x)$ we mean $u(x)+v(x)$.)

In other words, the derivative of the sum of two functions is just the sum of their derivatives. We'll now prove that this is true for any pair of functions u and v, provided that those functions have derivatives. Since we don't know in advance what functions u and v are, we can't use any specific information about the functions or the slopes of their graphs; all we have to work with is the formal definition of the derivative.

When we apply the definition of the derivative to the function $(u+v)(x)$ we get:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{(u+v)(x+\Delta x) - (u+v)(x)}{\Delta x}$$

Since $(u+v)(x)$ is just $u(x)+v(x)$,

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{u(x+\Delta x) + v(x+\Delta x) - u(x) - v(x)}{\Delta x}.$$

Combining like terms, we see that:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{u(x+\Delta x) - u(x) + v(x+\Delta x) - v(x)}{\Delta x}$$

or:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \left\{ \frac{u(x+\Delta x) - u(x)}{\Delta x} + \frac{v(x+\Delta x) - v(x)}{\Delta x} \right\}.$$

Because u and v are differentiable (and therefore continuous), the limit of the sum is the sum of the limits. Therefore:

$$(u+v)'(x) = \lim_{\Delta x \to 0} \frac{u(x+\Delta x) - u(x)}{\Delta x} + \lim_{\Delta x \to 0} \frac{v(x+\Delta x) - v(x)}{\Delta x}.$$

The two limits above match the definition of the derivatives of u and v, so we've shown that $(u+v)'(x) = u'(x) + v'(x)$.

Derivative of $\sin x$, Algebraic Proof

A specific derivative formula tells us how to take the derivative of a specific function: if $f(x) = x^n$ then $f'(x) = nx^{n-1}$. We'll now compute a specific formula for the derivative of the function $\sin x$.

As before, we begin with the definition of the derivative:

$$\frac{d}{dx}\sin x = \lim_{\Delta x \to 0} \frac{\sin(x + \Delta x) - \sin(x)}{\Delta x}$$

You may remember the following angle sum formula from high school:

$$\sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)$$

This lets us untangle the x from the Δx as follows:

$$\frac{d}{dx}\sin x = \lim_{\Delta x \to 0} \frac{\sin x \cos \Delta x + \cos x \sin \Delta x - \sin(x)}{\Delta x}.$$

We can simplify this expression using some basic algebraic facts:

$$\begin{aligned}
\frac{d}{dx}\sin x &= \lim_{\Delta x \to 0} \left[\frac{\sin x \cos \Delta x - \sin x}{\Delta x} + \frac{\cos x \sin \Delta x}{\Delta x}\right] \\
&= \lim_{\Delta x \to 0} \left[\frac{\sin x(\cos \Delta x - 1)}{\Delta x} + \frac{\cos x \sin \Delta x}{\Delta x}\right] \\
&= \lim_{\Delta x \to 0} \left[\sin x \left(\frac{\cos \Delta x - 1}{\Delta x}\right) + \cos x \left(\frac{\sin \Delta x}{\Delta x}\right)\right] \\
\frac{d}{dx}\sin x &= \lim_{\Delta x \to 0} \sin x \left(\frac{\cos \Delta x - 1}{\Delta x}\right) + \lim_{\Delta x \to 0} \cos x \left(\frac{\sin \Delta x}{\Delta x}\right)
\end{aligned}$$

We now have two familiar functions – $\sin x$ and $\cos x$ – and two ugly looking fractions to deal with. The fractions may be familiar from our discussion of removable discontinuities.

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{\cos \Delta x - 1}{\Delta x} &= 0 \\
\lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} &= 1.
\end{aligned}$$

Using these (as yet unproven) facts,

$$\lim_{\Delta x \to 0} \sin x \left(\frac{\cos \Delta x - 1}{\Delta x}\right) + \lim_{\Delta x \to 0} \cos x \left(\frac{\sin \Delta x}{\Delta x}\right)$$

simplifies to $\sin x \cdot 0 + \cos x \cdot 1 = \cos x$. We conclude:

$$\frac{d}{dx}\sin x = \cos x$$

Derivative of $\cos x$.

What is the specific formula for the derivative of the function $\cos x$?

This calculation is very similar to that of the derivative of $\sin(x)$. If you get stuck on a step here it may help to go back and review the corresponding step there.

As in the calculation of $\frac{d}{dx}\sin x$, we begin with the definition of the derivative:

$$\frac{d}{dx}\cos x = \lim_{\Delta x \to 0} \frac{\cos(x + \Delta x) - \cos(x)}{\Delta x}$$

Use the angle sum formula $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$ and then simplify:

$$\begin{aligned}
\frac{d}{dx}\cos x &= \lim_{\Delta x \to 0} \frac{\cos x \cos \Delta x - \sin x \sin \Delta x - \cos x}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left[\frac{\cos x \cos \Delta x - \cos x}{\Delta x} + \frac{-\sin x \sin \Delta x}{\Delta x} \right] \\
&= \lim_{\Delta x \to 0} \left[\frac{\cos x (\cos \Delta x - 1)}{\Delta x} + \frac{-\sin x \sin \Delta x}{\Delta x} \right] \\
&= \lim_{\Delta x \to 0} \left[\cos x \left(\frac{\cos \Delta x - 1}{\Delta x} \right) + (-\sin x) \left(\frac{\sin \Delta x}{\Delta x} \right) \right] \\
\frac{d}{dx}\cos x &= \lim_{\Delta x \to 0} \cos x \left(\frac{\cos \Delta x - 1}{\Delta x} \right) + \lim_{\Delta x \to 0} (-\sin x) \left(\frac{\sin \Delta x}{\Delta x} \right)
\end{aligned}$$

Once again we use the following (unproven) facts:

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{\cos \Delta x - 1}{\Delta x} &= 0 \text{ (A)} \\
\lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} &= 1. \text{ (B)}
\end{aligned}$$

And we conclude:

$$\begin{aligned}
\frac{d}{dx}\cos x &= \lim_{\Delta x \to 0} \cos x \left(\frac{\cos \Delta x - 1}{\Delta x} \right) + \lim_{\Delta x \to 0} (-\sin x) \left(\frac{\sin \Delta x}{\Delta x} \right) \\
&= \cos x \cdot 0 + (-\sin x) \cdot 1 \\
\frac{d}{dx}\cos x &= -sin(x).
\end{aligned}$$

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

In order to compute specific formulas for the derivatives of $\sin(x)$ and $\cos(x)$, we needed to understand the behavior of $\sin(x)/x$ near $x = 0$ (property B). In his lecture, Professor Jerison uses the definition of $\sin(\theta)$ as the y-coordinate of a point on the unit circle to prove that $\lim_{\theta\to 0}(\sin(\theta)/\theta) = 1$.

We switch from using x to using θ because we want to start thinking about the sine function as describing a ratio of sides in the triangle shown in Figure 1. The variable we're interested in is an angle, not a horizontal position, so we discuss $\sin(\theta)/\theta$ rather than $\sin(x)/x$.

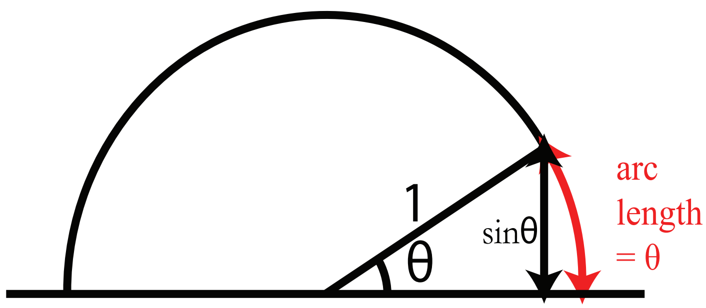


Figure 1: A circle of radius 1 with an arc of angle θ.

Our argument depends on the fact that when the radius of the circle shown in Figure 1 is 1, θ is the length of the highlighted arc. This is true when the angle θ is described in radians but NOT when it is measured in degrees.

Also, since the radius of the circle is 1, $\sin(\theta) = \frac{|\text{opposite}|}{|\text{hypotenuse}|}$ equals the length of the edge indicated (the hypotenuse has length 1).

In other words, $\sin(\theta)/\theta$ is the ratio of edge length to arc length. When $\theta = \pi/2$ rad, $\sin(\theta) = 1$ and $\sin(\theta)/\theta = 2/\pi \cong 2/3$. When $\theta = \pi/4$ rad, $\sin(\theta) = \sqrt{2}/2$ and $\sin(\theta)/\theta = 2\sqrt{2}/\pi \cong 9/10$. What will happen to the value of $\sin(\theta)/\theta$ as the value of θ gets closer and closer to 0 radians?

We see from Figure 2 that as θ shrinks, the length $\sin(\theta)$ of the segment gets closer and closer to the length θ of the curved arc. We conclude that as $\theta \to 0$,

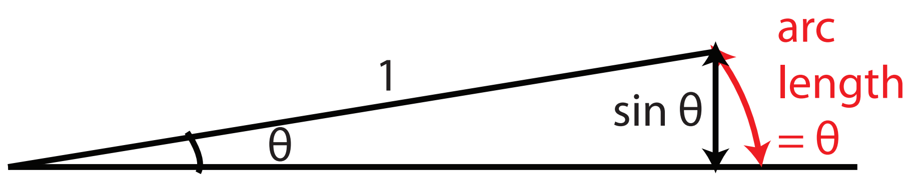


Figure 2: The sector in Fig. 1 as θ becomes very small

$\dfrac{\sin\theta}{\theta} \to 1$. In other words,

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1.$$

This technique of comparing very short segments of curves to straight line segments is a powerful and important one in calculus; it is used several times in this lecture.

$$\lim_{\theta \to 0} \frac{1 - \cos\theta}{\theta} = 0$$

While calculating the derivatives of $\cos(x)$ and $\sin(x)$, Professor Jerison said that $\lim_{\theta \to 0} \frac{1 - \cos\theta}{\theta} = 0$. This is true, but in order to be certain that our derivative formulas are correct we should understand *why* it's true.

As in the discussion of $\sin(\theta)/\theta$, our explanation involves looking at a diagram of the unit circle and comparing an arc with length θ to a straight line segment. (Remember that θ is measured in radians!) As shown in Figure 1, the vertical distance between the endpoints of the arc is $\cos\theta$, and the horizontal distance between the ends of the arc is $1 - \cos\theta$.

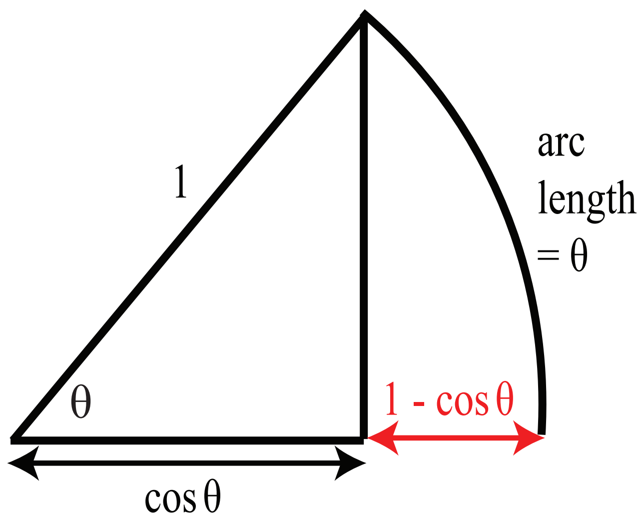


Figure 1: Same figure as for $\frac{\sin x}{x}$ except that the horizontal distance between the edge of the triangle and the perimeter of the circle is marked

From Fig. 2 we can see that as $\theta \to 0$, the horizontal distance $1 - \cos\theta$ between endpoints of the arc (what Professor Jerison calls "the gap") gets much smaller than the length θ of the arc. Hence, $\frac{1 - \cos\theta}{\theta} \to 0$.

If you find this hard to believe it may be helpful to use a calculator to verify that if x is small, $1 - \cos x$ is much smaller. You might also study the graph of $y = 1 - \cos x$ near $x = 0$ or use a web application to compare the distance $1 - \cos\theta$ to the arc length θ for very small angles θ.

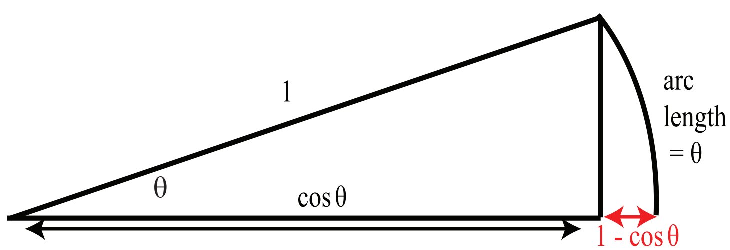


Figure 2: The sector in Fig. 1 as θ becomes very small

Questions and Answers

It's difficult to visualize the relationship between the arc length θ and the segment length $1 - \cos(\theta)$ in the geometric argument that:

$$\lim_{x \to 0} \frac{1 - \cos(\theta)}{\theta} = 0.$$

Professor Jerison spends ten minutes answering student questions and clarifying his argument.

Doesn't θ also tend to 0?

Yes. Whenever we take a derivative we're dividing by a quantity that tends to zero. The quantity $\lim_{x \to 0} \frac{1-\cos(\theta)}{\theta}$ happens to be the derivative of $\cos\theta$ evaluated at zero. It should not surprise us that this ratio tends to zero divided by zero. What we're interested in is that as θ goes to zero the straight line approximating the arc with length θ becomes much longer than the "gap" of length $1 - \cos\theta$.

As the angle θ decreases and we "zoom in", it turns out that the ratio between the size of the "gap" and the length of the "bowstring" gets smaller and smaller – the gap shrinks faster than the bowstring. So, as the measure of angle θ approaches 0, the ratio $\frac{1-\cos(\theta)}{\theta}$ also approaches 0.

In the example of $\dfrac{\sin\theta}{\theta}$ we were comparing two quantities that were about the same size and which approached zero at approximately the same rate, so their ratio approached 1.

Where is the Stata Center?

The Stata Center is a building at MIT that is approximately 100 meters away from the hall in which this lecture is given.

Is $\frac{\cos\theta - 1}{\theta}$ the same as $\frac{1-\cos(\theta)}{\theta}$?

$\dfrac{\cos\theta - 1}{\theta}$ is the negative of $\dfrac{1 - \cos(\theta)}{\theta}$.

$$-\frac{1 - \cos(\theta)}{\theta} = \frac{-1 + \cos\theta}{\theta} = \frac{\cos\theta - 1}{\theta}$$

What do you mean by "arc length"?

The arc in question is the portion of the unit circle swept out by the angle θ. Its arc length is its length.

Why is its length θ?

Because we are measuring the angle θ in radians and because the radius of the circle has length 1, the length of the arc swept out by θ will equal the angle θ.

This is especially important in the discussion of $\frac{\sin\theta}{\theta}$. If we were measuring θ in degrees, for example, we would have to compute the arc length before we could compare it to the vertical distance $\sin\theta$. The arc length only equals θ when θ is measured in radians.

The best way to measure θ in this calculation is by length along the unit circle, which is what radians are. Unfortunately, we won't see a proof of this until near the end of the course.

Why is the length of the radius equal to 1?

When working with trig functions like $\cos(x)$ it's extremely useful to think about a circle with radius 1 centered at the point $(0, 0)$ because a point (x, y) at angle θ on the circle will have coordinates $(\cos\theta, \sin\theta)$.

The radius shown on the blackboard looks like it's more than one unit long because we "zoomed in" on it as θ became too small to see.

I'm having trouble visualizing how $\frac{1-\cos(\theta)}{\theta} \to 0$.

You are probably not the only one.

First, visualize angle θ getting smaller and smaller.

When θ is very, very small, we can't see what's going on unless we "blow up" or "zoom in on" the picture. (In fact, if we let θ decrease all the way to 0, the triangle collapses and $\frac{1-\cos(\theta)}{\theta} = \frac{0}{0}$. That's why we have to use a limit.)

To estimate the value of $\frac{1-\cos(\theta)}{\theta}$ as θ shrinks, we need to look at geometric interpretations of both the numerator $(1 - \cos\theta)$ and the denominator θ and compare them.

The numerator is the length of the tiny "gap" and the denominator is half the length of the "bow string". From the picture, we see that the gap is much shorter than the bow string when θ is small.

As θ shrinks, we zoom in further, the bow of the circular arc gets closer and closer to the bow string, and so the gap gets smaller and smaller while the bow string appears to stay the same length. So the ratio of the size of the gap to the length of the bow string gets smaller and smaller, approaching zero.

Not only does the size of the gap go to zero, it goes to zero faster than the length of the bow string does.

A geometric proof that the derivative of $\sin x$ is $\cos x$.

At the start of the lecture we saw an algebraic proof that the derivative of $\sin x$ is $\cos x$. While this proof was perfectly valid, it was somewhat abstract – it did not make use of the definition of the sine function.

The proof that $\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1$ did use the unit circle definition of the sine of an angle. It also showed that when $x = 0$ the derivative of $\sin x$ is 1:

$$\begin{aligned} \frac{d}{dx} \sin x|_{x=0} &= \lim_{\Delta x \to 0} \frac{\sin(0 + \Delta x) - \sin 0}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{\sin \Delta x - 0}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} \\ &= 1. \end{aligned}$$

We'll now prove that the derivative of $\sin \theta$ is $\cos \theta$ directly from the definition of the sine function as the ratio $\frac{|\text{opposite}|}{|\text{hypotenuse}|}$ of the side lengths of a right triangle.

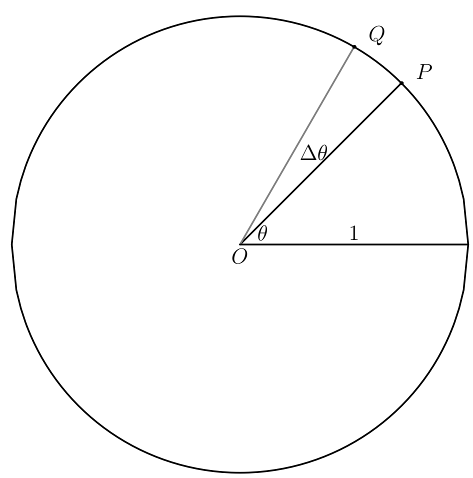


Figure 1: Point P has vertical position $\sin \theta$.

We start with a point P on the unit circle centered at O and the angle θ associated with P. As indicated in Figure 1, $\sin \theta$ is the vertical distance between P and the x-axis. Next, we add a small amount $\Delta\theta$ to angle θ; let Q be the point on the unit circle at angle $\theta + \Delta\theta$. The y-coordinate of Q is $\sin(\theta + \Delta\theta)$. To find the rate of change of $\sin \theta$ with respect to θ we just need to find the rate of change of $y = \sin \theta$.

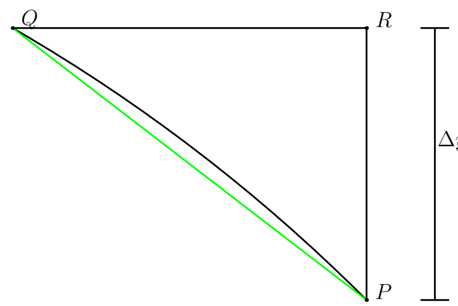


Figure 2: When $\Delta\theta$ is small, $\overparen{PQ} \approx \overline{PQ}$. Find $\frac{dy}{d\theta}$.

As shown in Figure 2, $\Delta y = |PR|$ and segment PQ is a straight line approximation of the circular arc PQ. If $\Delta\theta$ is small enough, segment PQ and arc PQ are practically the same, so $|PQ| \approx \Delta\theta$.

We're trying to find Δy. Since we know the length of the hypotenuse PQ, all we need is the measure of $\angle QPR$ to solve for $\Delta y = |PR|$.

Since $\Delta\theta$ is small, segment PQ is (nearly) tangent to the circle, and so angle $\angle OPQ$ is (nearly) a right angle. We know that PR is vertical, we know that θ is the angle OP makes with the horizontal, and we can combine these facts to prove that $\angle RPQ$ and θ are (nearly) congruent angles.[1]

The arc length $\Delta\theta$ is approximately equal to the length $|PR|$ of the hypotenuse and angle RPQ is approximately equal to θ. By the definition of the cosine function we get $\cos\theta \approx \frac{|PR|}{\Delta\theta}$. But $|PR|$ is just the vertical distance between Q and P, which is just the difference between $\sin(\theta + \Delta\theta)$ and $\sin\theta$. In other words, when $\Delta\theta$ is very small,

$$\cos\theta \approx \frac{\sin(\theta + \Delta\theta) - \sin\theta}{\Delta\theta}.$$

As $\Delta\theta$ approaches 0, segment QP gets closer and closer to arc QP and angle QPO gets closer and closer to a right angle, so the value of $\frac{(\sin(\theta+\Delta\theta)-\sin\theta)}{\Delta\theta}$ gets closer and closer to $\cos\theta$. We conclude that:

$$\lim_{\Delta\theta\to 0} \frac{\sin(\theta + \Delta\theta) - \sin\theta}{\Delta\theta} = \cos\theta$$

and thus that the derivative of $\sin\theta$ is $\cos\theta$.

[1] Professor Jerison does this by rotating and translating angle θ to coincide with angle RPQ. Another way to see this is to extend segment RP until it intersects the horizontal line through O at point S, then note that $m\angle RPQ + m\angle QPO + m\angle OPS = \pi$ and also $\theta + m\angle PSO + m\angle OPS = \pi$. Since $m\angle QPO \cong m\angle PSO$, we get $m\angle RPQ \cong \theta$. (If $\theta > \pi/2$ a different, but similar, argument applies.)

General Derivative Rules

We've just seen some specific rules for taking the derivatives of the cosine and sine functions. Here are some general rules which we'll discuss in more detail later.

Product Rule

$$(uv)' = u'v + uv'$$

The way that you should remember this is by thinking about changing one function (u or v) at a time. This is a good general procedure when taking derivatives involving multiple functions.

Quotient Rule

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} \quad (v \neq 0).$$

You'll see proofs of these soon, and you should be able to prove facts like these for your homework and exams.

Introduction to General Rules for Differentiation

We've now seen several specific rules for differentiation; for example, x^n is nx^{n-1}. We've even seen a few examples using this formula. We've also seen some general rules for extending these calculations. For instance, $(cu)' = c \cdot u'$ and $(u + v)' = u' + v'$.

Today we'll learn more general rules; how to differentiate a product of functions, a quotient of functions, and best of all a composition of functions. At the end we'll learn something about higher derivatives.

Product formula (General)

The product rule tells us how to take the derivative of the product of two functions:

$$(uv)' = u'v + uv'$$

This seems odd — that the product of the derivatives is a sum, rather than just a product of derivatives — but in a minute we'll see why this happens.

First, we'll use this rule to take the derivative of the product $x^n \sin x$ – a function we would not be able to differentiate without this rule. Here the first function, u is x^n and the second function v is $\sin x$. According to the specific rule for the derivative of x^n, the derivative u' must be nx^{n-1}. If $v = \sin x$ then $v' = \cos x$. The product rule tells us that $(uv)' = u'v + uv'$, so

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

By applying this rule repeatedly, we can find derivatives of more complicated products:

$$\begin{aligned}
(uvw)' &= u'(vw) + u(vw)' \\
&= u'vw + u(v'w + vw') \\
&= u'vw + uv'w + uvw'.
\end{aligned}$$

Now let's see why this is true:

$$\begin{aligned}
(uv)' &= \lim_{\Delta x \to 0} \frac{(uv)(x + \Delta x) - (uv)(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{u(x + \Delta x)v(x + \Delta x) - u(x)v(x)}{\Delta x}
\end{aligned}$$

We want our final formula to appear in terms of u, v, u' and v'. We know that $u' = \lim_{\Delta x \to 0} \frac{u(x + \Delta x) - u(x)}{\Delta x}$, and we see that $u(x + \Delta x)v(x + \Delta x) - u(x)v(x)$ looks a little bit like $(u(x + \Delta x) - u(x))v(x)$. By using a little bit of algebra we can get $(u(x + \Delta x) - u(x))v(x)$ to appear in our formula; this process is described below.

First, notice that:

$$u(x + \Delta x)v(x) - u(x + \Delta x)v(x) = 0.$$

Adding zero to the numerator doesn't change the value of our expression, so:

$$(uv)' = \lim_{\Delta x \to 0} \frac{u(x + \Delta x)v(x) - u(x)v(x) + u(x + \Delta x)v(x + \Delta x) - u(x + \Delta x)v(x)}{\Delta x}.$$

We then re-arrange that expression to get:

$$(uv)' = \lim_{\Delta x \to 0} \left[\left(\frac{u(x + \Delta x) - u(x)}{\Delta x} \right) v(x) + u(x + \Delta x) \left(\frac{v(x + \Delta x) - v(x)}{\Delta x} \right) \right]$$

We proved that if u and v are differentiable they must be continuous, so the limit of the sum is the sum of the limits:

$$\left[\lim_{\Delta x \to 0} \frac{u(x+\Delta x) - u(x)}{\Delta x}\right] v(x) + \lim_{\Delta x \to 0} \left(u(x+\Delta x) \left[\frac{v(x+\Delta x) - v(x)}{\Delta x}\right]\right)$$

or in other words,

$$(uv)' = u'(x)v(x) + u(x)v'(x).$$

Note: we also used the fact that:

$$\lim_{\Delta x \to 0} u(x+\Delta x) = u(x),$$

which is true because u is differentiable and therefore continuous.

Quotient Rule

Now that we know the product rule we can find the derivatives of many more functions than we used to be able to. Our next step toward "differentiating everything" will be to learn a formula for differentiating quotients (fractions). The rule is:

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

Why is this true? The definition of the derivative tells us that:

$$\left(\frac{u}{v}\right)' = \lim_{\Delta x \to 0} \frac{\dfrac{u(x+\Delta x)}{v(x+\Delta x)} - \dfrac{u(x)}{v(x)}}{\Delta x}$$

This is an unwieldy expression. We'll start to make sense of it by simplifying the numerator and by creating two new variables $\Delta u = u(x+\Delta x) - u(x)$ and $\Delta v = v(x+\Delta x) - v(x)$.

$$\begin{aligned}
\frac{u(x+\Delta x)}{v(x+\Delta x)} - \frac{u(x)}{v(x)} &= \frac{u+\Delta u}{v+\Delta v} - \frac{u}{v} \quad (\text{since } u(x+\Delta x) = u(x) + \Delta u) \\
&= \frac{(u+\Delta u)v - u(v+\Delta v)}{(v+\Delta v)v} \quad (\text{common denominator}) \\
&= \frac{uv + (\Delta u)v - uv + u(\Delta v)}{(v+\Delta v)v} \quad (\text{distribute } u \text{ and } v) \\
&= \frac{(\Delta u)v - u(\Delta v)}{(v+\Delta v)v} \quad (\text{because } uv - uv = 0)
\end{aligned}$$

Now that we've simplified the numerator, we can use it to simplify the difference quotient:

$$\begin{aligned}
\frac{\dfrac{u(x+\Delta x)}{v(x+\Delta x)} - \dfrac{u(x)}{v(x)}}{\Delta x} &= \frac{\dfrac{(\Delta u)v - u(\Delta v)}{(v+\Delta v)v}}{\Delta x} \\
&= \frac{1}{\Delta x}\,\frac{(\Delta u)v - u(\Delta v)}{(v+\Delta v)v} \\
&= \frac{\left(\dfrac{\Delta u}{\Delta x}\right)v - u\left(\dfrac{\Delta v}{\Delta x}\right)}{(v+\Delta v)v}
\end{aligned}$$

we're assuming that v is differentiable and therefore continuous, so $\lim\limits_{x\to 0} v(x+\Delta x) = v(x)$. Hence, by the definition of the derivative,

$$\frac{\left(\dfrac{\Delta u}{\Delta x}\right)v - u\left(\dfrac{\Delta v}{\Delta x}\right)}{(v+\Delta v)v} \longrightarrow \frac{v\left(\dfrac{du}{dx}\right) - u\left(\dfrac{dv}{dx}\right)}{v^2} \qquad \text{as } \Delta x \to 0.$$

We conclude that:

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}.$$

Example: Reciprocals

Let's use the quotient rule in a simple example. The quotient rule tells us that:

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{\frac{du}{dx}v - u\frac{dv}{dx}}{v^2}$$

In this example u will be 1, so we'll be finding the derivative of $\frac{1}{v}$, the reciprocal of v.

$$\frac{d}{dx}\left(\frac{1}{v}\right) = ?$$

We're going to use the formula above. We know $u = 1$ and $v = v$, so we still need to find $\frac{du}{dx}$ and $\frac{dv}{dx}$ before we can apply the formula.

The derivative of a constant (like 1) is zero, so $\frac{du}{dx} = 0$. We don't know what v is, so we'll just write $\frac{dv}{dx} = v'$. Plugging all this in to the quotient rule formula we get:

$$\begin{aligned} \frac{d}{dx}\left(\frac{1}{v}\right) &= \frac{0 \cdot v - 1v'}{v^2} \\ &= \frac{-v'}{v^2} \\ &= -v^{-2}v' \end{aligned}$$

Now we have a general formula that lets us differentiate reciprocals! Next, let's use this formula to see what happens when $u = 1$ and $v = x^n$. Here again $\frac{du}{dx} = 0$ and now $v' = \frac{d}{dx}x^n = nx^{n-1}$.

$$\begin{aligned} \frac{d}{dx}\left(\frac{1}{x^n}\right) &= -v^{-2}v' \\ &= -(x^n)^{-2}(nx^{n-1}) \\ &= -x^{-2n}(nx^{n-1}) \\ &= -nx^{-n-1} \end{aligned}$$

But $\frac{1}{x^n} = x^{-n}$, which is x to a power. We have a rule for taking the derivative of x to a positive power; how does that compare to our new rule for the derivative of x to a negative power?

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

This agrees with the formula $\frac{d}{dx}x^n = nx^{n-1}$, so the quotient rule confirms that our rule for taking the derivative of x^n works even when n is negative.

Chain Rule

The product rule tells us how to find the derivative of a product of functions like $f(x) \cdot g(x)$. The composition or "chain" rule tells us how to find the derivative of a composition of functions like $f(g(x))$. Composition of functions is about substitution – you substitute a value for x into the formula for g, then you substitute the result into the formula for f. An example of a composition of two functions is $y = (\sin t)^{10}$ (which is usually written as $y = \sin^{10} t$).

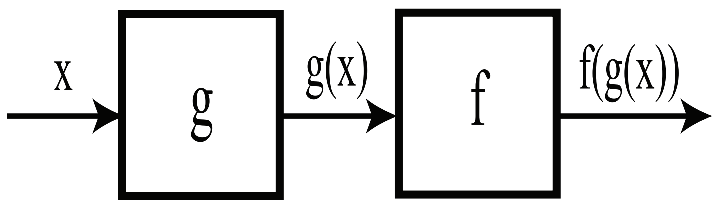


Figure 1: Composition of functions: $(f \circ g)(x) = f(g(x))$

One way to think about composition of functions is to use new variable names. For example, for the function $y = \sin^{10} t$ we can say $x = \sin t$ and then $y = x^{10}$. Notice that if you plug $x = \sin t$ in to the formula for y you get back to $y = \sin^{10} t$. It's good practice to introduce new variables when they're convenient, and this is one place where it's very convenient.

So, how do we find the derivative of a composition of functions? We're trying to find the slope of a tangent line; to do this we take a limit of slopes $\frac{\Delta y}{\Delta t}$ of secant lines. Here y is a function of x, x is a function of t, and we want to know how y changes with respect to the original variable t. Here again using that intermediate variable x is a big help:

$$\frac{\Delta y}{\Delta t} \quad = \quad \frac{\Delta y}{\Delta x} \cdot \frac{\Delta x}{\Delta t}$$

because when we perform the multiplication, the small change Δx cancels.

The derivative of y with respect to t is $\lim\limits_{\Delta t \to 0} \dfrac{\Delta y}{\Delta t}$; what happens when Δt gets small? Because $x = \sin t$ is a continuous function, as Δt approaches 0, Δx also approaches zero. It turns out that:

$$\lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} = \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} \qquad \longleftarrow \textbf{The Chain Rule!}$$

The derivative of a composition of functions is a product. In the example $y = (\sin t)^{10}$, we have the "inside function" $x = \sin t$ and the "outside function" $y = x^{10}$. The chain rule tells us to take the derivative of y with respect to x and multiply it by the derivative of x with respect to t.

The derivative of $y = x^{10}$ is $\dfrac{dy}{dx} = 10x^9$. The derivative of $x = \sin t$ is $\dfrac{dx}{dt} = \cos t$. The chain rule tells us that $\frac{d}{dt} \sin^{10} t = 10x^9 \cdot \cos t$. This is correct, but if a friend asked you for the derivative of $\sin^{10} t$ and you answered $10x^9 \cdot \cos t$ your friend wouldn't know what x stood for. The last step in this process is to rewrite x in terms of t:

$$\frac{d}{dt} \sin^{10} t = 10(\sin t)^9 \cdot \cos t = 10 \sin^9 t \cdot \cos t.$$

Here is another way of writing the chain rule:

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

Example: $\sin(10t)$

The chain rule (composition rule) says that $\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$. In other words, the derivative of the composition of functions $f(g(t))$ is the derivative of the outside function $f(x)$ times the derivative of the inside function $g(t)$.

For the example $\frac{d}{dt}\sin(10t)$, the inside function is $x = 10t$ and the outside function is $y = \sin x$. Using the rules we know, we can compute that $\frac{dy}{dx} = \cos x$ and $\frac{dx}{dt} = 10$, so:

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = \cos x \cdot 10.$$

Since we're the only ones who know the value of x in this formula, we replace x by $10t$ to get:

$$\frac{dy}{dt} = \cos(10t) \cdot 10 = 10\cos(10t).$$

Once you've had more practice using the chain rule, you won't always need the variable x that represents the inside function. When you look at $\frac{d}{dt}\sin(10t)$ you might say to yourself: "The derivative of the outside function, sine, is cosine. I'm plugging $10t$ into it. And the derivative of $10t$ is just 10. So $\frac{d}{dt}\sin(10t) = \cos(10t) \cdot 10$.

Higher Derivatives

Higher derivatives are derivatives of derivatives. Given a differentiable function $u = u(x)$ its derivative u' is a new function, which we may be able to differentiate again to get $(u')' = u''$.

For example, if $u(x) = \sin x$ then $u' = \cos x$ and $u'' = -\sin x$. We can go on: $(u'')' = u''' = -\cos x$ ($u''' = u^{(3)}$ is called the third derivative of u and u'' is the second derivative) and $u'''' = u^{(4)} = \sin x$. The function $\sin x$ is a special example – we won't usually "come back to" the function we started with.

Since there's more than one way to write derivatives, there's more than one notation for higher derivatives.

Notations

$f'(x)$	Df	$\frac{df}{dx}$	$\frac{d}{dx}f$
$f''(x)$	D^2f	$\frac{d^2f}{dx^2}$	$\left(\frac{d}{dx}\right)^2 f$
$f'''(x)$	D^3f	$\frac{d^3f}{dx^3}$	$\left(\frac{d}{dx}\right)^3 f$
$f^{(n)}(x)$	D^nf	$\frac{d^nf}{dx^n}$	$\left(\frac{d}{dx}\right)^{\mathrm{n}} f$

The symbols D and $\frac{d}{dx}$ represent "operators" which can be applied to a function. When you apply one of these operators to a function you get the derivative of that function.

Example: $D^n x^n$

Let's calculate the n^{th} derivative of x^n

$$D^n x^n = ? \quad (n = 1, 2, 3, ...)$$

Let's start small and look for a pattern:

$$\begin{aligned}
Dx^n &= nx^{n-1} \\
D^2x^n &= n(n-1)x^{n-2} \\
D^3x^n &= n(n-1)(n-2)x^{n-3} \\
&\vdots \\
D^{n-1}x^n &= (n(n-1)(n-2)\cdots 2)x^1
\end{aligned}$$

We can guess this $(n-1)^{st}$ derivative from the pattern established by the first three derivatives. The power of x decreases by 1 at every step, so the power of x on the $(n-1)^{st}$ step will be 1. At each step we multiply the derivative by the power of x from the previous step, so at the $(n-1)^{st}$ step we'll be multiplying by the previous power 2 of x.

Differentiating one more time we get:

$$D^n x^n = (n(n-1)(n-2)\cdots 2 \cdot 1)1$$

The number $(n(n-1)(n-2)\cdots 2\cdot 1)$ is written $n!$ and is called "n factorial". What we've just seen forms the basis of a proof by mathematical induction that $D^n x^n = n!$. So $D^n x^n$ is a constant!

The final question for the lecture is: what is $D^{n+1}x^n$?

Answer: It's the derivative of a constant, so it's 0.

Introduction to Implicit Differentiation

We've learned a few specific and several general formulas for finding derivatives. Today we'll use the chain rule to further expand our ability to differentiate functions. Today's topic is implicit differentiation, which will allow us to differentiate many functions we haven't been able to differentiate yet.

Implicit Differentiation (Rational Exponent Rule)

We know that if n is an integer then the derivative of $y = x^n$ is nx^{n-1}. Does this formula still work if n is not an integer? I.e. is it true that:

$$\frac{d}{dx}(x^a) = ax^{a-1}.$$

We proved this formula using the definition of the derivative and the binomial theorem for $a = 1, 2, ...$ From this, we also got the formula for $a = -1, -2, ...$ Now we'll extend this formula to cover rational numbers $a = \frac{m}{n}$ as well. In particular, this will let us take the derivative of $y = \sqrt[n]{x} = x^{1/n}$.

Suppose $y = x^{\frac{m}{n}}$, where m and n are integers. We want to compute $\frac{dy}{dx}$. None of the rules we've learned so far seem helpful here, and if we use the definition of the derivative we'll get stuck trying to simplify $(x - \Delta x)^{m/n}$. We need a new idea.

The thing that's keeping us from using the definition of the derivative is that the denominator of n in the exponent forces us to take the n^{th} root of x. We could solve this problem by raising both sides of the equation to the n^{th} power:

$$\begin{aligned} y &= x^{\frac{m}{n}} \\ y^n &= (x^{\frac{m}{n}})^n \\ y^n &= x^{\frac{m}{n}\cdot n} \\ y^n &= x^m \end{aligned}$$

What happens if we try to take the derivative now by applying the operator $\frac{d}{dx}$? We have a rule for finding the derivative of a variable raised to an integer power; we can use this rule on both sides of the equation $y^n = x^m$.

$$\begin{aligned} y^n &= x^m \\ \frac{d}{dx}y^n &= \frac{d}{dx}x^m \end{aligned}$$

How do we compute $\frac{d}{dx}y^n$? We know that y is a function of x, so we can apply the chain rule with outside function y^n and inside function y. Suppose $u = y^n$. Then the chain rule tells us:

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

So

$$\frac{d}{dx}y^n = \left(\frac{d}{dy}y^n\right)\frac{dy}{dx} = ny^{n-1}\frac{dy}{dx}.$$

On the right hand side of the equation we have $\frac{d}{dx}x^m = mx^{m-1}$, so we end up with:

$$\begin{aligned}\frac{d}{dx}y^n &= \frac{d}{dx}x^m \\ ny^{n-1}\frac{dy}{dx} = mx^{m-1}\end{aligned}$$

We're left with only one unknown quantity in this equation — $\frac{dy}{dx}$ — which is exactly what we're trying to find. Can we solve for $\dfrac{dy}{dx}$ and use this to find the derivative of $y = x^{m/n}$? We can, but we need to use a lot of algebra to do it.

By dividing both sides by ny^{n-1} we get:

$$\frac{dy}{dx} = \frac{m}{n}\frac{x^{m-1}}{y^{n-1}}$$

This looks promising but we want our answer in terms of x, without any y's mixed in. To get rid of the y we can now substitute $x^{m/n}$ for y. (We couldn't have done this before taking the derivative because we don't know how to take the derivative of $x^{m/n}$ — that's the whole point!)

$$\begin{aligned}\frac{dy}{dx} &= \frac{m}{n}\left(\frac{x^{m-1}}{y^{n-1}}\right) \\ &= \frac{m}{n}\left(\frac{x^{m-1}}{(x^{m/n})^{(n-1)}}\right) \\ &= \frac{m}{n}\left(\frac{x^{m-1}}{x^{(m/n)\cdot(n-1)}}\right) \\ &= \frac{m}{n}\frac{x^{m-1}}{x^{m(n-1)/n}} \\ &= \frac{m}{n}x^{((m-1)-\ \frac{m(n-1)}{n})} \\ &= \frac{m}{n}x^{\frac{n(m-1)}{n}\ -\ \frac{m(n-1)}{n}} \\ &= \frac{m}{n}x^{\frac{n(m-1)-m(n-1)}{n}} \\ &= \frac{m}{n}x^{\frac{nm-n-nm+m}{n}} \\ &= \frac{m}{n}x^{\frac{m-n}{n}} \\ &= \frac{m}{n}x^{(\frac{m}{n}-\frac{n}{n})} \\ \text{So,}\quad \frac{dy}{dx} &= \frac{m}{n}x^{(\frac{m}{n}-1)}\end{aligned}$$

This is the answer we were hoping to get! We now know that for any rational number a, the derivative of x^a is ax^{a-1}.

Slope of a line tangent to a circle – direct version

A circle of radius 1 centered at the origin consists of all points (x, y) for which $x^2 + y^2 = 1$. This equation does not describe a function of x (i.e. it cannot be written in the form $y = f(x)$). Indeed, any vertical line drawn through the interior of the circle meets the circle in two points — every x has two corresponding y values. Let's see what goes wrong if we attempt to solve the equation of a circle for y in terms of x.

$$\begin{aligned} x^2 + y^2 &= 1 \\ x^2 + y^2 - x^2 &= 1 - x^2 \\ y^2 &= 1 - x^2 \\ y &= \pm\sqrt{1 - x^2} \end{aligned}$$

This still isn't a function because we get two choices for y — positive or negative. However, we do get a function if we look just at the positive case (i.e. at just the top half of the circle), and we can then find $\frac{dy}{dx}$, which will be the slope of a line tangent to the top half of the circle.

To compute this derivative, we first convert the square root into a fractional exponent so that we can use the rule from the previous example.

$$y = \sqrt{1 - x^2} = (1 - x^2)^{\frac{1}{2}}$$

Next, we need to use the chain rule to differentiate $y = (1 - x^2)^{\frac{1}{2}}$. The outside function is $u^{1/2}$ and the inside function is $1 - x^2$, so the chain rule tells us that

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

$$\frac{dy}{dx} = \frac{1}{2}u^{-1/2} \cdot (-2x) = -x \cdot (1 - x^2)^{-1/2} = \frac{-x}{\sqrt{1 - x^2}}.$$

If we want, we can use the fact that $y = \sqrt{1 - x^2}$ to rewrite this as $y' = -x/y$.

We conclude that the slope of the line tangent to a point (x, y) on the top half of the unit circle is $-x/y$.

Slope of a line tangent to a circle – implicit version

We just finished calculating the slope of the line tangent to a point (x, y) on the top half of the unit circle. In this calculation we started by solving the equation $x^2 + y^2 = 1$ for y, chose one "branch" of the solution to work with, then used the chain rule, the power rule and some algebra of exponents to compute the derivative $\frac{dy}{dx} = -\frac{x}{y}$.

We'll now see how we could have used implicit differentiation to do the same calculation much more easily. In fact, we'll find the slope of a line tangent to *any* point on the unit circle.

We don't need to solve for y — we can just apply the operator $\frac{d}{dx}$ to both sides of the original equation:

$$\begin{aligned} x^2 + y^2 &= 1 \\ \frac{d}{dx}\left(x^2 + y^2\right) &= \frac{d}{dx}(1) \\ \frac{d}{dx}(x^2) + \frac{d}{dx}(y^2) &= 0 \end{aligned}$$

We can easily take the derivative of the first term. For the second term, applying the chain rule with the inside function y and outside function u^2 gives us:

$$\begin{aligned} 2x + 2y\frac{dy}{dx} &= 0 \\ 2y\frac{dy}{dx} &= -2x \\ \frac{dy}{dx} &= -\frac{x}{y} \end{aligned}$$

This is the same answer, but we didn't have to restrict ourselves to just the top half of the circle or use any square roots. Implicit differentiation made this calculation much easier.

Implicit Differentiation Example

How would we find $y' = \frac{dy}{dx}$ if $y^4 + xy^2 - 2 = 0$?

We could use a trick to solve this explicitly — think of the above equation as a quadratic equation in the variable y^2 then apply the quadratic formula:

$$\begin{aligned} y^2 &= \frac{-x \pm \sqrt{x^2+8}}{2}, \\ &\text{so} \\ y &= \pm\sqrt{\frac{-x \pm \sqrt{x^2+8}}{2}}. \end{aligned}$$

Since we see $\pm$ twice in this equation, there are four possible branches to consider. This means that to be thorough we'd want to compute four different derivatives. This is a lot of work.

Instead, we can compute $\frac{dy}{dx}$ using implicit differentiation. As always, we start by applying $\frac{d}{dx}$ to both sides:

$$\begin{aligned} \frac{d}{dx}(y^4 + xy^2 - 2) &= \frac{d}{dx}0 \\ \frac{d}{dx}(y^4) + \frac{d}{dx}(xy^2) - \frac{d}{dx}2 &= 0 \\ 4y^3\frac{dy}{dx} + (y^2 + x \cdot 2y\frac{dy}{dx}) - 0 &= 0 \\ 4y^3\frac{dy}{dx} + 2xy\frac{dy}{dx} &= -y^2 \\ (4y^3 + 2xy)\frac{dy}{dx} &= -y^2 \\ \frac{dy}{dx} &= \frac{-y^2}{4y^3 + 2xy} \end{aligned}$$

In lecture Professor Jerison used the shorthand y' for the derivative; here we use $\frac{dy}{dx}$ to make it clear that we are differentiating with respect to x.

Derivative of the Inverse of a Function

One very important application of implicit differentiation is to finding derivatives of inverse functions.

We start with a simple example. We might simplify the equation $y = \sqrt{x}$ ($x > 0$) by squaring both sides to get $y^2 = x$. We could use function notation here to say that $y = f(x) = \sqrt{x}$ and $x = g(y) = y^2$.

In general, we look for functions $y = f(x)$ and $g(y) = x$ for which $g(f(x)) = x$. If this is the case, then g is the inverse of f (we write $g = f^{-1}$) and f is the inverse of g (we write $f = g^{-1}$).

How are the graphs of a function and its inverse related? We start by graphing $f(x) = \sqrt{x}$. Next we want to graph the inverse of f, which is $g(y) = x$. But this is exactly the graph we just drew. To compare the graphs of the functions f and f^{-1} we have to exchange x and y in the equation for f^{-1}. So to compare $f(x) = \sqrt{x}$ to its inverse we replace y's by x's and graph $g(x) = x^2$.

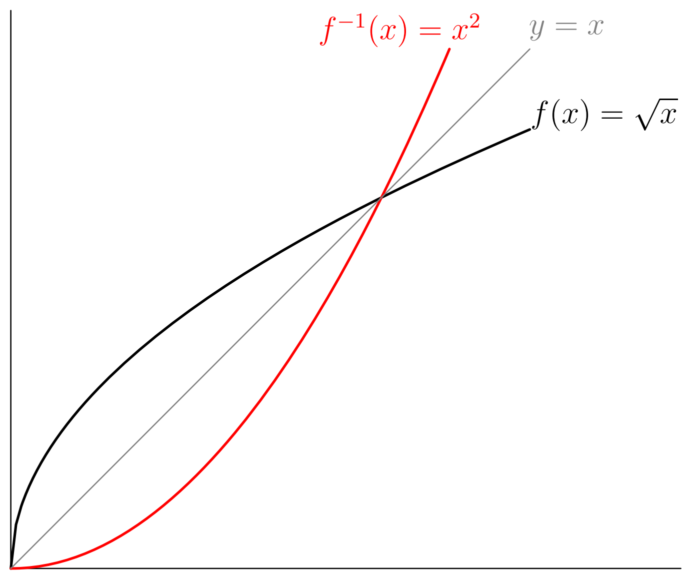


Figure 1: The graph of f^{-1} is the reflection of the graph of f across the line $y = x$

In general, if you have the graph of a function f you can find the graph of f^{-1} by exchanging the x- and y-coordinates of all the points on the graph. In other words, the graph of f^{-1} is the reflection of the graph of f across the line $y = x$.

This suggests that if $\frac{dy}{dx}$ is the slope of a line tangent to the graph of f, then

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

is the slope of a line tangent to the graph of f^{-1}. We could use the definition of the derivative and properties of inverse functions to turn this suggestion into a proof, but it's easier to prove using implicit differentiation.

Let's use implicit differentiation to find the derivative of the inverse function:

$$\begin{aligned} y &= f(x) \\ f^{-1}(y) &= x \\ \frac{d}{dx}(f^{-1}(y)) &= \frac{d}{dx}(x) = 1 \end{aligned}$$

By the chain rule:

$$\frac{d}{dy}(f^{-1}(y))\frac{dy}{dx} = 1$$

so

$$\frac{d}{dy}(f^{-1}(y)) = \frac{1}{\frac{dy}{dx}}.$$

Implicit differentiation allows us to find the derivative of the inverse function $x = f^{-1}(y)$ whenever we know the derivative of the original function $y = f(x)$.

Derivative of $\arctan(x)$

Let's use our formula for the derivative of an inverse function to find the derivative of the inverse of the tangent function: $y = \tan^{-1} x = \arctan x$.

We simplify the equation by taking the tangent of both sides:

$$\begin{aligned} y &= \tan^{-1} x \\ \tan y &= \tan(\tan^{-1} x) \\ \tan y &= x \end{aligned}$$

To get an idea what to expect, we start by graphing the tangent function (see Figure 1). The function $\tan(x)$ is defined for $-\frac{\pi}{2} < x < \frac{\pi}{2}$. It's graph extends from negative infinity to positive infinity.

If we reflect the graph of $\tan x$ across the line $y = x$ we get the graph of $y = \arctan x$ (Figure 2). Note that the function $\arctan x$ is defined for all values of x from minus infinity to infinity, and $\lim_{x\to\infty} \tan^{-1} x = \frac{\pi}{2}$.

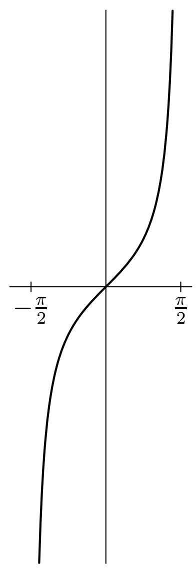


Figure 1: Graph of the tangent function.

You may know that:

$$\begin{aligned} \frac{d}{dy} \tan y &= \frac{d}{dy} \frac{\sin y}{\cos y} \\ &\vdots \\ &= \frac{1}{\cos^2 y} \\ &= \sec^2 y \end{aligned}$$

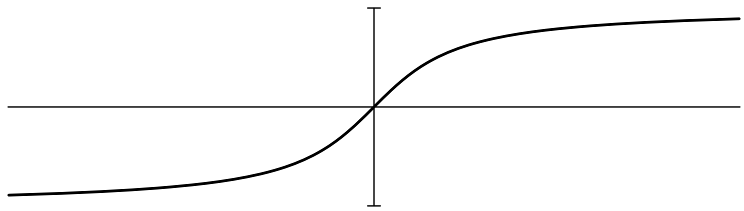


Figure 2: Graph of $\tan^{-1} x$.

(If you haven't seen this before, it's good exercise to use the quotient rule to verify it!)

We can now use implicit differentiation to take the derivative of both sides of our original equation to get:

$$\begin{aligned}
\tan y &= x \\
\frac{d}{dx}(\tan(y)) &= \frac{d}{dx}x \\
\text{(Chain Rule)} \quad \frac{d}{dy}(\tan(y))\frac{dy}{dx} &= 1 \\
\left(\frac{1}{\cos^2(y)}\right)\frac{dy}{dx} &= 1 \\
\frac{dy}{dx} &= \cos^2(y)
\end{aligned}$$

Or equivalently, $y' = \cos^2 y$. Unfortunately, we want the derivative as a function of x, not of y. We must now plug in the original formula for y, which was $y = \tan^{-1} x$, to get $y' = \cos^2(\arctan(x))$. This is a correct answer but it can be simplified tremendously. We'll use some geometry to simplify it.

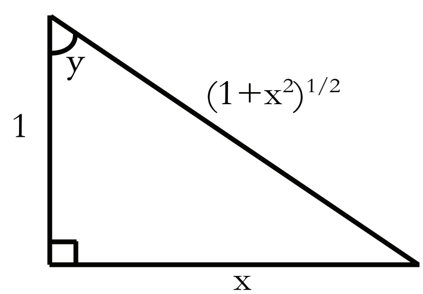


Figure 3: Triangle with angles and lengths corresponding to those in the example.

In this triangle, $\tan(y) = x$ so $y = \arctan(x)$. The Pythagorean theorem tells us the length of the hypotenuse:

$$h = \sqrt{1 + x^2}$$

and we can now compute:

$$\cos(y) = \frac{1}{\sqrt{1 + x^2}}.$$

From this, we get:

$$\cos^2(y) = \left(\frac{1}{\sqrt{1 + x^2}}\right)^2 = \frac{1}{1 + x^2}$$

so:

$$\frac{dy}{dx} = \frac{1}{1 + x^2}.$$

In other words,

$$\frac{d}{dx} \arctan(x) = \frac{1}{1 + x^2}.$$

Derivative of $\arcsin(x)$

For a final example, we quickly find the derivative of $y = \sin^{-1} x = \arcsin x$.

As usual, we simplify the equation by taking the sine of both sides:

$$\begin{aligned} y &= \sin^{-1} x \\ \sin y &= x \end{aligned}$$

We next take the derivative of both sides of the equation and solve for $y' = \frac{dy}{dx}$.

$$\begin{aligned} \sin y &= x \\ (\cos y) \cdot y' &= 1 \\ y' &= \frac{1}{\cos y} \end{aligned}$$

We want to rewrite this in terms of $x = \sin y$. Luckily there is a simple trig. identity relating $\cos y$ to $\sin y$. We can solve it for $\cos y$ and "plug in".

$$\begin{aligned} \cos^2 y + \sin^2 y &= 1 \\ \cos^2 y &= 1 - \sin^2 y \\ \cos y &= \sqrt{1 - \sin^2 y} \quad (\cos y > 0 \text{ on the range of } y = \sin^{-1} x) \end{aligned}$$

Plugging this in to our equation for $y' = \frac{d}{dx} \sin^{-1} x$ we get:

$$y' = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - \sin^2 y}} = \frac{1}{\sqrt{1 - x^2}}.$$

Notice that we made a choice between a positive and negative square root when solving for $\cos y$. We chose the positive square root because we usually define $\sin^{-1} x$ to have outputs between $-\pi/2$ and $\pi/2$, and the cosine function is always positive on this interval.

When dealing with inverse functions we are often faced with choices like this; when in doubt draw a graph and be sure your choices make sense in the context of your problem.

Differentiating Logs and Exponentials

We finally discuss derivatives of exponential and logarithmic functions. These are probably the only functions you're aware of that you're still unable to differentiate. Logs and exponentials are as fundamental as trigonometric functions, if not more so.

Working with exponents

We start out with "base" number a. This number a must be positive, and we're going to assume $a > 1$ to make it easier to draw graphs.

What is the derivative of a^x? We'll start to answer this question by reviewing what we know about exponents.

To begin with, we know that:

$$a^0 = 1; \quad a^1 = a; \quad a^2 = a \cdot a; \quad a^3 = a \cdot a \cdot a \quad \ldots$$

In general,

$$a^{x_1+x_2} \quad = \quad a^{x_1} a^{x_2}$$

Together with the first two properties, this describes the exponential function a^x.

From these properties, we can derive:

$$(a^{x_1})^{x_2} \; = \; a^{x_1 x_2}$$

and we can easily evaluate a^n for any positive integer n. For negative integers, we can see from the fact that $a^m \cdot a^{-m} = a^{m-m} = 1$ that $a^{-m} = \frac{1}{a^m}$.

We want to be able to evaluate a^x for any number x; not just for integers. We start by defining a^x for rational values of x:

$$a^{\frac{p}{q}} \; = \; \sqrt[q]{a^p} \quad \text{(where } p \text{ and } q \text{ are integers.)}$$

Since $a^{1/2} \cdot a^{1/2} = a^1 = \sqrt{a} \cdot \sqrt{a}$, this seems like a reasonable definition.

All that's left is to define a^x for irrational numbers; we do this by "filling in" the gaps in the function to make it continuous. This is what your calculator does when you ask it for the value of $3^{\sqrt{2}}$ or 2^π. It can't give you an exact answer, so it gives you a decimal (rational) number very close to the exact answer.

Take some time and sketch the graph of 2^x to "get a feel" for how exponential functions work.

a^x and the Definition of the Derivative

Our goal is to calculate the derivative $\frac{d}{dx}a^x$. It's going to take us a while.

We start by writing down the definition of the derivative

$$\frac{d}{dx}a^x = \lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x}$$

We can use the rule $a^{x_1+x_2} = a^{x_1}a^{x_2}$ to factor out a^x:

$$\begin{aligned} \frac{d}{dx}a^x &= \lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{a^x a^{\Delta x} - a^x}{\Delta x} \\ &= \lim_{\Delta x \to 0} a^x \frac{a^{\Delta x} - 1}{\Delta x} \end{aligned}$$

As we're taking this limit, we're holding a and x fixed while Δx changes (approaches zero). This means that for the purposes of taking this limit, a^x is a constant. We can therefore factor the constant multiple out of the limit to get:

$$\frac{d}{dx}a^x = a^x \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}$$

We've made a good start at finding the derivative of a^x; let's look at what we have so far. We can see from our calculations that $\frac{d}{dx}a^x$ is a^x times some multiple whose value we don't yet know. Let's call that multiple $M(a)$:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

Using this definition of $M(a)$, we can say that $\frac{d}{dx}a^x = M(a)a^x$.

Slope of the tangent to a^x

We defined a function $M(a)$ as follows:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

This definition allows us to say that $\frac{d}{dx} a^x = M(a) a^x$. In order to understand the derivative of a^x we must understand $M(a)$; we next look at two different ways of thinking about $M(a)$.

First, if we plug $x = 0$ in to the definition of the derivative of a^x we get:

$$\begin{aligned} \left. \frac{d}{dx} a^x \right|_{x=0} &= \lim_{\Delta x \to 0} \left. \frac{a^{x+\Delta x} - a^x}{\Delta x} \right|_{x=0} \\ &= \lim_{\Delta x \to 0} \frac{a^{0+\Delta x} - a^0}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x} \\ &= M(a) \end{aligned}$$

(or we could simply observe that $\frac{d}{dx} a^x|_{x=0} = M(a)a^0 = M(a)$). So $M(a)$ is the value of the derivative of a^x when $x = 0$.

Remember that the derivative tells us the slope of the tangent line to the graph. So $M(a)$ can also be thought of as the slope of the graph of $y = a^x$ at $x = 0$.

Note that the shape of the graph of a^x depends on the choice of a, so for different values for a we'll get different tangent lines and different values for $M(a)$.

Because $\frac{d}{dx} a^x = M(a) a^x$, we only need to know the slope of the line tangent to the graph at $x = 0$ in order to figure out the slope of the tangent line at any point on the graph!

Remember that when we computed the derivative of the sine function we worked hard to compute the value of $\lim_{x \to 0} \frac{\sin x}{x}$. This value is just the derivative of $\sin x$ when $x = 0$ — check this yourself by writing down the definition of the derivative of $\sin x$ and replacing x by 0. In order to get a general formula for the derivative of the sine function we first had to know the value of its derivative when $x = 0$.

The formula for $a^{x+\Delta x}$ is simpler than the one for $\sin(x + \Delta x)$, so the first part of our calculation of $\frac{d}{dx} a^x$ was easier than the corresponding calculation for $\sin x$. But when we try to compute:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}$$

we get stuck. We were able to use radians and the unit circle to find $\lim_{x \to 0} \frac{\sin x}{x}$,

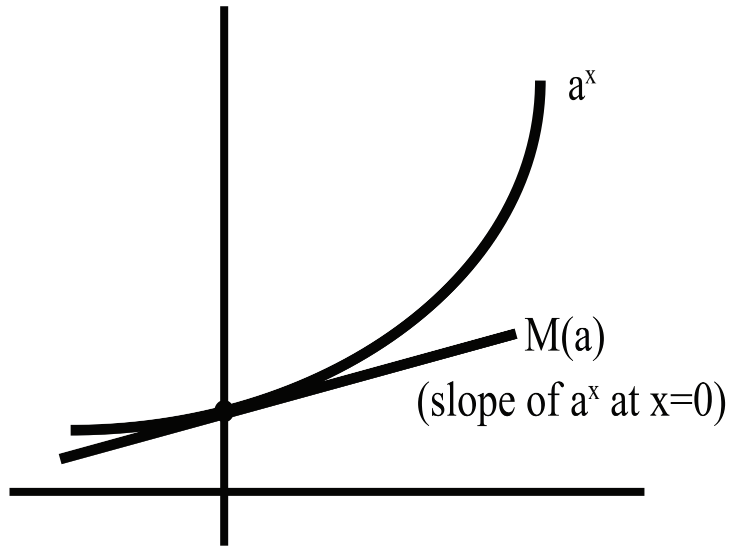


Figure 1: Geometric definition of $M(a)$

but we don't have a good way to find the exact slope of the tangent line to $y = a^x$ at $x = 0$.

Definition of e

Recall that:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

is the value for which $\frac{d}{dx}a^x = M(a)a^x$, the value of the derivative of a^x when $x = 0$, and the slope of the graph of $y = a^x$ at $x = 0$. We need to know what $M(a)$ is in order to find out what the derivative of a^x is. It turns out that the easiest way to understand $M(a)$ is to give up trying to calculate it and to *define* e as the number such that $M(e) = 1$.

Leaving aside the question of whether such a number e exists, let's discuss what such a number would do for us. Since $M(e) = 1$,

$$\frac{d}{dx}e^x = e^x.$$

This is an incredibly important formula and is the only thing we've said so far this lecture that you need to memorize. Also, the slope of the tangent line to $y = e^x$ at $x = 0$ has slope 1. You can confirm this by plugging $x = 0$ into $\frac{d}{dx}e^x = e^x$.

But we still don't know what e is, or even if there is such a number. How do we know that there is *any* number a for which the slope of the tangent line to $y = a^x$ is 1 when $x = 0$?

First notice that as the base a increases, the graph of $y = a^x$ gets steeper. Is the slope ever 1?

If $a = 1$, $a^x = 1$ for all x and the slope of the tangent line to the (very simple) graph at $x = 0$ is 0. Although we may not be able to compute the slope exactly, we can use secant lines to estimate the slope $M(a)$ for $a = 2$ and $a = 4$ geometrically. Look at the graph of 2^x in Fig. 1. The secant line from $(0, 1)$ to $(1, 2)$ of the graph $y = 2^x$ has slope 1. We can see from the picture that the slope of $y = 2^x$ at $x = 0$ is less than the slope of this secant line: $M(2) < 1$ (see Fig. 1).

Next, look at the graph of 4^x in Fig. 2. The secant line from $(-\frac{1}{2}, \frac{1}{2})$ to $(1, 0)$ on the graph of $y = 4^x$ has slope 1. We see that the slope of $y = 4^x$ at $x = 0$ is greater than the slope of the secant, so $M(4) > 1$ (see Fig. 2).

Assuming our function M is continuous, we conclude that somewhere in between 2 and 4 there is a base whose slope at $x = 0$ is 1.

Thus we can *define* e to be the unique number such that

$$M(e) = 1$$

or, to put it another way,

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1$$

or, to put it yet another way,

$$\frac{d}{dx}(e^x) = 1 \quad \text{at } x = 0$$

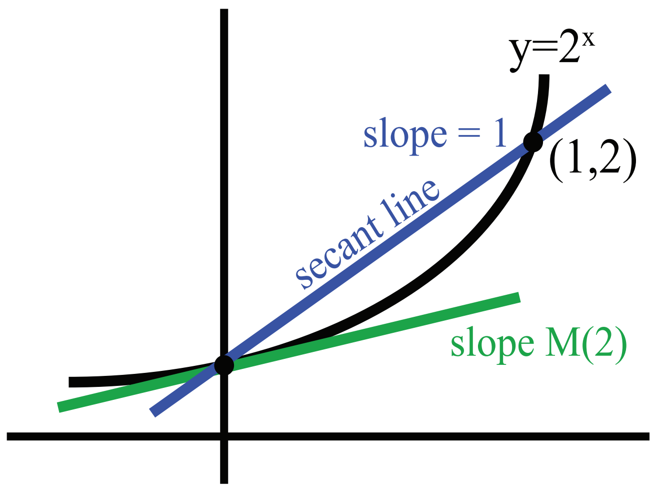


Figure 1: Slope $M(2) < 1$

Another way to convince ourselves that e must exist is to start with the graph of $f(x) = 2^x$ (recalling that $M(2) < 1$) and think about the function $f(kx) = 2^{kx}$. As k increases, the graph of $y = f(kx)$ is compressed horizontally and the slope of the tangent line to the graph of $y = f(x)$ continuously grows steeper. So, for some value of k between 1 and 2, the slope of that tangent line must be 1. So e exists and is between $2^1 = 2$ and $2^2 = 4$.

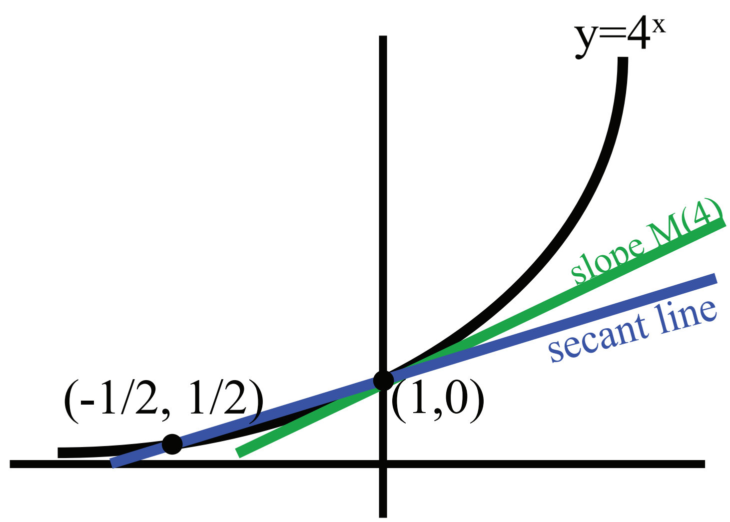


Figure 2: Slope $M(4) > 1$

Natural log (inverse function of e^x)

Recall that:

$$M(a) = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

is the value for which $\frac{d}{dx} a^x = M(a)a^x$, the value of the derivative of a^x when $x = 0$, and the slope of the graph of $y = a^x$ at $x = 0$. To understand $M(a)$ better, we study the natural log function $\ln(x)$, which is the inverse of the function e^x. This function is defined as follows:

$$\text{If } \ y = e^x, \ \text{ then } \ \ln(y) = x$$

or

$$\text{If } \ w = \ln(x), \ \text{ then } \ e^w = x$$

Before we go any further, let's review some properties of this function:

$$\ln(x_1 x_2) = \ln x_1 + \ln x_2$$

$$\ln 1 = 0$$

$$\ln e = 1$$

These can be derived from the definition of $\ln x$ as the inverse of the function e^x, the definition of e, and the rules of exponents we reviewed at the start of lecture.

We can also figure out what the graph of $\ln x$ must look like. We know roughly what the graph of e^x looks like, and the graph of $\ln x$ is just the reflection of that graph across the line $y = x$. Try sketching the graph of $\ln x$ yourself.

You should notice the following important facts about the graph of $\ln x$. Since e^x is always positive, the domain (set of possible inputs) of $\ln x$ includes only the positive numbers. The entire graph of $\ln x$ lies to the right of the y-axis. Since $e^0 = 1$, $\ln 1 = 0$ and the graph of $\ln x$ goes through the point $(1, 0)$. And finally, since the slope of the tangent line to $y = e^x$ is 1 where the graph crosses the y-axis, the slope of the graph of $y = \ln x$ must be $1/1 = 1$ where the graph crosses the x-axis.

We know that $\frac{d}{dx} e^x = e^x$. To find $\frac{d}{dx} \ln x$ we'll use implicit differentiation as we did in previous lectures.

We start with $w = \ln(x)$ and compute $\dfrac{dw}{dx} = \dfrac{d}{dx} \ln x$. We don't have a good way to do this directly, but since $w = \ln(x)$, we know $e^w = e^{\ln(x)} = x$. We now use implicit differentiation to take the derivative of both sides of this equation.

$$\begin{aligned} \frac{d}{dx}(e^w) &= \frac{d}{dx}(x) \\ \frac{d}{dw}(e^w)\frac{dw}{dx} &= 1 \end{aligned}$$

$$
\begin{aligned}
e^w \frac{dw}{dx} &= 1 \\
\frac{dw}{dx} &= \frac{1}{e^w} = \frac{1}{x}
\end{aligned}
$$

So

$$\frac{d}{dx}(\ln(x)) = \frac{1}{x}$$

This is another formula worth memorizing.

$\frac{d}{dx}a^x$?

We now want to learn to differentiate *any* exponential a^x. There are two roughly equivalent methods we can use:

Method 1: Convert a^x to something with base e and use the chain rule.

Because $\ln x$ is the inverse function to e^x we can rewrite a as $e^{\ln(a)}$. Thus:

$$a^x = \left(e^{\ln(a)}\right)^x = e^{x\ln(a)}$$

That looks like it might be tricky to differentiate. Let's work up to it:

$$\begin{aligned} \frac{d}{dx}e^x &= e^x \\ \frac{d}{dx}e^{3x} &= 3e^{3x} \quad \text{(by the chain rule)} \end{aligned}$$

Remember, $\ln(a)$ is just a constant like 3, not a variable. Therefore:

$$\frac{d}{dx}e^{(\ln a)x} = (\ln a)e^{(\ln a)x}$$

or

$$\frac{d}{dx}a^x = (\ln a)a^x$$

This is a common type of calculation; you should practice it until you are comfortable with it. You may either memorize formulas for $\frac{d}{dx}e^{kx}$ and $\frac{d}{dx}a^x$ or re-derive them every time you need them.

Recall that $\frac{d}{dx}a^x = M(a)\cdot a^x$. So finally we know the value of $M(a)$:

$$M(a) = \ln(a)$$

The Most Natural Logarithmic Function

At times in your life you might find yourself tempted to use logarithmic functions with bases other than e; for example, $\log_2 n$ or $\log_{10} r$. We claim that $\ln x$, the natural logarithm or log base e, is the most natural choice of logarithmic function.

Today's example is from the field of economics. Imagine that the price of a stock you own goes down by a dollar; how does that affect you?

That depends on a lot of things. In particular, it depends on whether the original stock price was a dollar or a hundred dollars.

As another example, at the time this lecture was given the London Exchange FTSE index had closed down 27.9 points. This statistic is almost meaningless unless you know the actual total of the index, which was 6,432 at the time.

A more meaningful statistic is the change in the price divided by the price:

$$\frac{\Delta p}{p} = \frac{27.9}{6432} \approx 0.43\%.$$

This tells us that the FTSE index dropped by 0.43% of its total value today.

A day is a relatively short time in the life of an investment so if price p is a function of time t with $\Delta t = 1$ day,

$$\frac{\Delta p}{1 \text{ day}} \approx \frac{dp}{dt} = p'.$$

So instead of looking at $\frac{\Delta p}{p}$ we could discuss $\frac{p'}{p}$, which is just the derivative of the natural log of p.

$$\frac{p'}{p} = (\ln p)'$$

This is the formula for logarithmic differentiation, and it is used all the time by economists and people who model prices of things.

There's no point in using log base ten or log base two, because when you take the derivative of those functions you get an extra constant factor in the denominator:

$$(\log_{10} p)' = \frac{p'}{p \ln 10}.$$

This is just one example — any variable that has to do with ratios is going to involve logarithms. We'll see more of this when we study applications of derivatives. You may have guessed this already; the derivative of $\ln x$ is quite elegant and simple. The most elegant and simple ideas that are often the most powerful, as well.

The Functions 10^x and 2^x

We computed that $\frac{d}{dx}a^x = (\ln a)a^x$.

So

$$\frac{d}{dx}2^x = (\ln 2)2^x$$

and

$$\frac{d}{dx}10^x = (\ln 10)10^x.$$

Even if we insist on starting with another base, like 2 or 10, the natural logarithm appears. They come up naturally, independent of our human preferences like base 2 or base 10. The base e may seem strange at first, but it comes up everywhere. After a while you'll learn to appreciate just how natural it is.

$\dfrac{d}{dx}a^x$, part 2

We're learning to differentiate *any* exponential a^x. This is the second of two possible methods.

Method 2: Logarithmic differentiation

It turns out that sometimes it is hard to differentiate a function u and easier to differentiate $\ln u$ (for example, $u = e^{x^2+6}$.) We'd like to be able to use $\frac{d}{dx}\ln u$ to find $\frac{d}{dx}u$.

The chain rule tells us that $\frac{d}{dx}\ln u = \frac{d\ln u}{du}\frac{du}{dx}$, and we know that $\frac{d}{du}\ln u = \frac{1}{u}\frac{du}{dx}$, so

$$(\ln u)' = u'/u.$$

How does this help us compute $\frac{d}{dx}a^x$?

$$\begin{aligned} u &= a^x \\ \ln u &= \ln(a^x) \\ \ln u &= x\ln a \end{aligned}$$

This is pretty easy to differentiate because $\ln a$ is a constant:

$$(\ln u)' = \ln a.$$

Since $(\ln u)' = u'/u$, $u' = u(\ln u)'$. So $\frac{d}{dx}a^x = a^x \ln a = (\ln a)a^x$.

This uses the same arithmetic as the first method, but we don't have to convert to base e.

Another Example of Logarithmic Differentiation

This example could be done equally well by converting to base e, but we're going to do it using logarithmic differentiation. Recall that the rule we use for logarithmic differentiation is $(\ln u)' = u'/u$.

Here we have a "moving" (non-constant) exponent and a moving base.

Example: Let $v = x^x$. Find v'.

First, we take the natural log of both sides to see that $\ln v = \ln(x^x) = x \ln x$.

Next, we differentiate both sides of the equation, using the product rule and the rule for the derivative of $\ln x$ on the right hand side:

$$(\ln v)' = \ln x + x \cdot \frac{1}{x}.$$

Now apply the formula $(\ln u)' = u'/u$. to get:

$$v'/v = 1 + \ln x$$

Plugging in x^x for v and solving for v', we get:

$$\begin{aligned} \frac{v'}{x^x} &= 1 + \ln x \\ v' &= x^x(1 + \ln x) \\ \frac{d}{dx}x^x &= x^x(1 + \ln x) \end{aligned}$$

The Power Rule

What is the derivative of $\frac{d}{dx}x^r$? We answered this question first for positive integer values of r, for all integers, and then for rational values of r:

$$\frac{d}{dx}x^r = rx^{r-1}$$

We'll now prove that this is true for any *real* number r. We can do this two ways:

1st method: base e

Since $x = e^{\ln x}$, we can say:

$$\begin{aligned} x^r &= \left(e^{\ln x}\right)^r \\ x^r &= e^{r \ln x} \end{aligned}$$

We take the derivative of both sides to get:

$$\begin{aligned} \frac{d}{dx}x^r &= \frac{d}{dx}e^{r\ln x} = e^{r\ln x}\frac{d}{dx}(r\ln x) \quad \text{(by the chain rule)} \\ &= e^{r\ln x}\left(\frac{r}{x}\right) \quad \text{(remember } r \text{ is constant)} \\ &= x^r\left(\frac{r}{x}\right) \quad \text{(because } x^r = e^{r\ln x}\text{)} \\ \frac{d}{dx}x^r &= rx^{r-1} \end{aligned}$$

2nd method: logarithmic differentiation

We define $f(x) = x^r$, and take the natural log of both sides to get $\ln f = r\ln x$. The technique of logarithmic differentiation requires us to we plug our function into the formula:

$$(\ln f)' = \frac{f'}{f}$$

So we first compute:

$$\begin{aligned} \ln f &= \ln x^r \\ \ln f &= r\ln x \end{aligned}$$

And then take the derivative of both sides to get:

$$(\ln f)' = \frac{r}{x}$$

Since $(\ln f)' = \frac{f'}{f}$, we have:

$$f' = f(\ln f)' \quad = \quad x^r \left(\frac{r}{x}\right) = rx^{r-1}.$$

Look over the two methods again – the calculations are almost the same. This is typical. To use the second method we had to introduce a new symbol like u or f. In the first method we had to deal with exponents. It's worthwhile know both methods.

Another Moving Exponent

Find the value of:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

Technically this is not a calculus problem, but we will use some calculus to solve it. There are two reasons to discuss this now – first that the answer is very interesting and second that it has another moving exponent – the exponent n in the problem is changing as we take the limit.

Whenever we're faced with a moving exponent our first step is to use a logarithm to turn the exponent into a multiple:

$$\ln\left[\left(1+\frac{1}{n}\right)^n\right] = n\ln\left(1+\frac{1}{n}\right).$$

Now we want to start thinking about the limit of this quantity as n approaches infinity. We've had a lot of practice thinking about limits as Δx approaches zero and very little practice with numbers approaching infinity, so it makes sense to try to rephrase this from a question about a very large number n to a question about a very small number Δx.

The quantity $\Delta x = 1/n$ will approach zero as n goes to infinity. If $\Delta x = 1/n$ then $n = 1/\Delta x$, and we get:

$$\lim_{n\to\infty}\left[n\ln\left(1+\frac{1}{n}\right)\right] = \lim_{\Delta x\to 0}\left[\frac{1}{\Delta x}\ln(1+\Delta x)\right].$$

This doesn't look like much of an improvement, but by subtracting $0 = \ln 1$ from $\ln(1+\Delta x)$ we can put it in a familiar form:

$$\begin{aligned}
\lim_{\Delta x\to 0}\left[\frac{1}{\Delta x}\ln(1+\Delta x)\right] &= \lim_{\Delta x\to 0}\left[\frac{1}{\Delta x}\left(\ln(1+\Delta x)-\ln 1\right)\right]\\
&= \lim_{\Delta x\to 0}\frac{\ln(1+\Delta x)-\ln 1}{\Delta x}\\
&= \left.\frac{d}{dx}\ln x\right|_{x=1}\\
&= \left.\frac{1}{x}\right|_{x=1}\\
&= 1
\end{aligned}$$

By strategically subtracting zero ($\ln 1$), we were able to turn this ugly limit into a difference equation, which we could then evaluate using calculus.

Now we just have to work backward to figure out the answer to our original question.

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty} e^{\ln\left[\left(1+\frac{1}{n}\right)^n\right]}$$

$$
\begin{aligned}
&= e^{\lim_{n\to\infty} \ln\left[\left(1+\frac{1}{n}\right)^n\right]} \\
&= e^1 \\
&= e
\end{aligned}
$$

That's right,

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$$

and we now have a way to get a numerical value for e. Using this formula we can find the value of e with as much precision as our calculators will allow. For example,

$$\left(1+\frac{1}{10000}\right)^{10000} \cong 2.7182.$$

A Formula for e

We calculated that if $a_k = (1 + \frac{1}{k})^k$, then $\lim_{k\to\infty} a_k = e$ by first showing that $\lim_{k\to\infty} \ln a_k = 1$. Since $e^{\ln a_k} = a_k$, as k goes to infinity $a_k = e^{\ln a_k}$ will tend toward $e^1 = e$.

The key point here was that $a_k = e^{\ln a_k}$; that the natural log function is the inverse of the exponential function.

Question: Shouldn't $\ln a_k$ tend towards zero, because a_k tends toward 1?

Answer: It's true that $(1+\frac{1}{k})$ tends toward 1, and so $\ln(1+\frac{1}{k})$ tends toward 0. But that's not the limit we want; we're asking about $\ln a_k = k \cdot \ln(1 + \frac{1}{k})$. As $\ln(1 + \frac{1}{k})$ is tending toward 0, k tends toward infinity. That's why we needed to use limits and derivatives to figure out what the limit of this expression was.

We know that $\lim_{k\to\infty} a_k = e$, and all equalities can be read in both directions. So $e = \lim_{k\to\infty} \left(1 + \frac{1}{k}\right)^k$. In other words, this limit is a formula for e. By looking at the formula from a different angle, we discover that we can use the expression $\left(1 + \frac{1}{k}\right)^k$ to compute a base e for which graph of e^x has slope 1 when $x = 0$.

Often we can improve our understanding of mathematics by looking at things in several different ways, and that's what we're going to be doing at the end of this lecture on derivatives.

Derivatives of Hyperbolic Sine and Cosine

Hyperbolic sine (pronounced "sinsh"):

$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$

Hyperbolic cosine (pronounced "cosh"):

$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$

$$\frac{d}{dx}\sinh(x) = \frac{d}{dx}\left(\frac{e^x - e^{-x}}{2}\right) = \frac{e^x - (-e^{-x})}{2} = \cosh(x)$$

Likewise,

$$\frac{d}{dx}\cosh(x) = \sinh(x)$$

(Note that this is different from $\frac{d}{dx}\cos(x)$.)

Important identity:

$$\cosh^2(x) - \sinh^2(x) = 1$$

Proof:

$$\begin{aligned}
\cosh^2(x) - \sinh^2(x) &= \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 \\
\cosh^2(x) - \sinh^2(x) &= \frac{1}{4}\left(e^{2x} + 2e^x e^{-x} + e^{-2x}\right) - \frac{1}{4}\left(e^{2x} - 2 + e^{-2x}\right) = \frac{1}{4}(2+2) = 1
\end{aligned}$$

Why are these functions called "hyperbolic"?
Let $u = \cosh(x)$ and $v = \sinh(x)$, then

$$u^2 - v^2 = 1$$

which is the equation of a hyperbola.

Regular trig functions are "circular" functions. If $u = \cos(x)$ and $v = \sin(x)$, then

$$u^2 + v^2 = 1$$

which is the equation of a circle.

Differentiation Formulas

General Differentiation Formulas

$$\begin{aligned}
(u+v)' &= u'+v' \\
(cu)' &= cu' \\
(uv)' &= u'v+uv' \quad \text{(product rule)} \\
\left(\frac{u}{v}\right)' &= \frac{u'v-uv'}{v^2} \quad \text{(quotient rule)} \\
\frac{d}{dx}f(u(x)) &= f'(u(x))\cdot u'(x) \quad \text{(chain rule)}
\end{aligned}$$

Implicit differentiation

Let's say you want to find y' from an equation like

$$y^3+3xy^2=8$$

Instead of solving for y and then taking its derivative, just take $\frac{d}{dx}$ of the whole thing. In this example,

$$\begin{aligned}
3y^2y'+6xyy'+3y^2 &= 0 \\
(3y^2+6xy)y' &= -3y^2 \\
y' &= \frac{-3y^2}{3y^2+6xy}
\end{aligned}$$

Note that this formula for y' involves both x and y.

As we see later in this lecture, implicit differentiation can be very useful for taking the derivatives of inverse functions and for logarithmic differentiation.

Specific differentiation formulas

You will be responsible for knowing formulas for the derivatives of these functions:

$$x^n,\ \sin^{-1}x,\ \tan^{-1}x,\ \sin x,\ \cos x,\ \tan x,\ \sec x,\ e^x,\ \ln x.$$

You may also be asked to derive formulas for the derivatives of these functions.

For example, let's calculate $\frac{d}{dx}\sec x$:

$$\frac{d}{dx}\sec x=\frac{d}{dx}\frac{1}{\cos x}=\frac{-(-\sin x)}{\cos^2 x}=\tan x\sec x$$

You may be asked to find $\frac{d}{dx}\sin x$ or $\frac{d}{dx}\cos x$ using the following information:

$$\begin{aligned}\lim_{h\to 0}\frac{\sin(h)}{h} &= 1\\ \lim_{h\to 0}\frac{\cos(h)-1}{h} &= 0\end{aligned}$$

The Chain Rule, Revisited

Why it's true

We didn't fully explain why the chain rule is true. We'll look at an example that should explain that. Consider the function

$$y = 10x + b.$$

Here y is changing ten times as fast as x, which is to say that $\frac{dy}{dx} = 10$.

Now, what if x is also a function of some variable t? If

$$x = 5t + a$$

then $\frac{dx}{dt} = 5$.

The chain rule says that if y is going ten times as fast as x, and x is going five times as fast as t, then y is going fifty times as fast as t. Algebraically, I replace x by $5t$ in the equation for y to get:

$$y = 10x + b = 10(5t + a) + b = 50t + 10a + b.$$

The consequence is that $\frac{dy}{dt} = 50 = 10 \cdot 5 = \frac{dy}{dx}\frac{dx}{dt}$. This is, in a nutshell, why the chain rule works and why these rates multiply.

Things it's good for

The chain rule can also make some of the other rules a little easier to remember or possibly to avoid. The messiest rule is perhaps the quotient rule. Notice that $\left(\frac{1}{v}\right)' = (v^{-1})'$. Instead of using the quotient rule here we can use the chain rule with the power -1 and the power law:

$$\left(\frac{1}{v}\right)' = (v^{-1})' = -v^{-2}v'.$$

Similarly,

$$\left(\frac{u}{v}\right)' = (uv^{-1})' = u'v^{-1} + u(-v^{-2})v'.$$

This explains the minus sign in the formula:

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}.$$

Derivative of secant

For example, let's use our formula for taking the derivative of $1/v$ to take the derivative of the secant function.

$$\frac{d}{dx}\sec x = \frac{d}{dx}\frac{1}{\cos x} = \frac{d}{dx}(\cos x)^{-1} = -(\cos x)^{-2}(-\sin x)$$

This is usually written in a different fashion; there are often many different ways of writing combinations of trigonometric functions. The standard way of writing this is:

$$\frac{d}{dx}\sec x = -(\cos x)^{-2}(-\sin x) = \frac{\sin x}{\cos^2 x} = \frac{1}{\cos x}\cdot\frac{\sin x}{\cos x} = \sec x \tan x.$$

This is the preferred form which uses the secant and tangent functions to avoid fractions and negative exponents. As you practice calculus with trigonometric functions you'll need to be aware of equivalent ways of describing the same result.

Derivative of $\ln(\sec x)$

Now let's use the chain rule to take the derivative of $\ln(\sec x)$).

$$\begin{aligned}
\frac{d}{dx}(\ln(\sec x)) &= \frac{(\sec x)'}{\sec x} \\
&= \frac{\sec x \tan x}{\sec x} \\
&= \tan x
\end{aligned}$$

Oddly enough, this strange looking function is not only interesting as a review of the chain rule. The natural log was invented before the exponential function by a man named Napier, exactly in order to evaluate functions like this.

People cared about these functions a lot because they were used in navigation. In order to quickly and accurately multiply sines and cosines together for navigation, Napier used a logarithm. Logarithms were invented long before people knew about exponents, and it was a surprise when it was discovered that they were connected to exponents. The natural log was invented before the log base ten and everything else, exactly for this kind of purpose.

Derivative of $(x^{10} + 8x)^6$

How do you take the derivative of $(x^{10} + 8x)^6$?

One thing you *don't* do is expand the expression by multiplying $(x^{10} + 8x)$ by itself repeatedly. Instead, we use the chain rule:

$$\frac{d}{dx}(x^{10} + 8x)^6 \quad = \quad 6(x^{10} + 8x)^5(10x^9 + 8)$$

Question: What form should we leave our answers in?

Answer: That's a very important question. If a computer answers a question with 500 million pages of printout, that answer is useless. If for some reason you need your answer in polynomial form you may have to expand $(x^{10} + 8x)^6$ after all. For the exam you may leave your answer in any form as long as it's correct. In particular, it's best not to try to simplify your answer unless you're specifically instructed to.

Derivative of $e^{x \tan^{-1} x}$

Finally, in the first lecture I promised you that you'd learn to differentiate *anything*— even something as complicated as

$$\frac{d}{dx} e^{x \tan^{-1} x}$$

So let's do it!

$$\frac{d}{dx} e^{uv} = e^{uv} \frac{d}{dx}(uv) = e^{uv}(u'v + uv')$$

Substituting,

$$\frac{d}{dx} e^{x \tan^{-1} x} = e^{x \tan^{-1} x} \left(\tan^{-1} x + x \left(\frac{1}{1 + x^2} \right) \right)$$

Exam 1 Review, Continued

The Definition of the Derivative

The main thing we talked about in the first part of the course was the definition of the derivative; one of our goals has been to understand its meaning. The formula for the derivative is:

$$\frac{d}{dx}f(x) = f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

This is a central focus of this course and you want to be able to recognize this formula in a number of forms.

How might you be asked to use this on a test? Our chief use of this formula occurred while finding specific formulas for derivatives. In fact, we used it to find many of the formulas you might be tested on:

$$x^n,\ \sin^{-1} x,\ \tan^{-1} x,\ \sin x,\ \cos x,\ \tan x,\ \sec x,\ e^x,\ \ln x.$$

Let's briefly review which functions we used this on, and what other facts we needed in finding our formulas. We used the definition of the derivative to compute formulas for the derivatives of

$$1/x,\ x^n,\ \sin x,\ \cos x,\ a^x,\ u \cdot v \text{ and } u/v.$$

To complete these calculations we needed to know the derivatives of $\sin x$, $\cos x$ and a^x at $x = 0$. To derive the product and quotient rules we needed to know the slopes of the individual functions u and v. It wasn't the definition of the derivative alone that got us these formulas, but in each case it got us to something simpler that we could use to get our formulas. That "getting to something simpler" is the sort of thing you should be able to do on a test.

But don't sit down and memorize the derivations of all those formulas to prepare for the test. What you need to learn is how to use the definition of the derivative to derive a formula for any function — there could be questions on e^x, $\frac{1}{x^2}$, or some other function.

Limits

There are also some fundamental limits you should know about for the test.

Recall that any equation can be read in two directions. In the case of the definition of the derivative, reading the equation backward tells us that if we know the slope of the graph of a function you can find the value of a limit:

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

For example, suppose you are asked to find:

$$\lim_{u \to 0} \frac{e^u - 1}{u}.$$

If you recognize that:

$$\lim_{u\to 0} \frac{e^u - 1}{u} = \frac{d}{du} e^u \bigg|_{u=0}$$

then it's easy to see that the answer is 1. The key is recognizing that

$$\lim_{u\to 0} \frac{e^u - 1}{u}$$

matches the formula

$$\lim_{\Delta x\to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

when $f(x) = e^x$, $\Delta x = u$, and $x = 0$.

Implicit Differentiation (continued)

For the test, you also should be able to use implicit differentiation to derive formulas for derivatives of functions like $\sin^{-1} x$ and $\ln x$. The power of implicit differentiation is that it lets you take a formula and simplify it as much as possible; you're not restricted to writing y as a function of x.

For instance, $\sin y = x$ is a much simpler equation than $y = \sin^{-1} x$. This second formula is easy to differentiate implicitly:

$$(\cos y)y' = 1$$

so

$$y' = \frac{1}{\cos y} = \frac{1}{\sqrt{1-x^2}}.$$

(If you're not sure why $\cos(\sin^{-1} x) = \sqrt{1-x^2}$, draw a right triangle whose "opposite" edge has length x and whose hypotenuse has length 1.)

Tangent Lines

Geometrically, a derivative is the slope of a tangent line to a graph. On the test you may be asked to compute the equation of a tangent line, to graph the function y' or to tell whether a function is differentiable by looking at its graph.

The first two types of questions are fairly straightforward, and we really only have one way to answer the third type of question. If a graph is differentiable, it must be possible to draw a line tangent to every point on the graph. This means that the limit of the secant lines must be well defined at every point on the graph. And this, in turn, means that the secant lines must have the same limits as they approach the tangent from the left and from the right.

One way to practice for the test is to graph a function like $y = \ln x$ and then try to draw the graph of y'. To do this, first draw a few tangent lines to the graph. Observe what you can about them – are their slopes all positive? All increasing? All the same? Do any tangent lines have slope zero? How steep is the steepest tangent line, or is there no steepest line? Then use that information to graph a curve whose height above the x-axis (approximately) equals the slopes of the tangents to the original graph.

Do not expect the graph of the derivative to look like the graph of the function. If possible, check your work by taking the derivative of the function whose graph you looked at.

Introduction to Linear Approximation

We're starting a new unit: applications of differentiation.

We're going to do two applications today. The first is linear approximations, which are encompassed by the single formula:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0),$$

but it will take at least half an hour to explain how this formula is used.

Linear Approximation to $\ln x$ at $x = 1$

If you have a curve $y = f(x)$, it is approximately the same as its tangent line $y = f(x_0) + f'(x_0)(x - x_0)$.

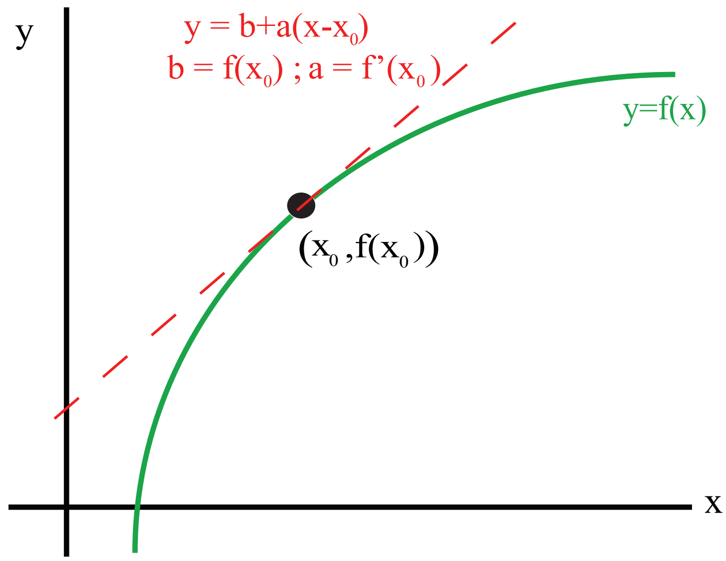


Figure 1: Tangent as a linear approximation to a curve

The tangent line approximates $f(x)$. It gives a good approximation near the tangent point x_0. As you move away from x_0, however, the approximation grows less accurate.

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

Example 1 Let $f(x) = \ln x$. Then $f'(x) = \frac{1}{x}$. We'll use the base point $x_0 = 1$ because we can easily evaluate $\ln 1 = 0$. Note also that $f'(x_0) = \frac{1}{1} = 1$.

Then the formula for linear approximation tells us that:

$$\begin{aligned} f(x) &\approx f(x_0) + f'(x_0)(x - x_0) \\ \ln x &\approx \ln(1) + 1(x - 1) \\ \ln x &\approx 0 + (x - 1) \\ \ln x &\approx (x - 1) \end{aligned}$$

Graph the curve $y = \ln x$ and the line $y = x - 1$. You'll see that the two graphs are very close together when $x = x_0 = 1$. You'll also see that they're only near each other when x is near 1.

The point of linear approximation is that the curve (in this case $y = \ln x$) is approximately the same as the tangent line ($y = x - 1$) when x is close to the base point x_0.

Linear Approximation and the Definition of the Derivative

Another way to understand the formula for linear approximation involves the definition of the derivative:

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

Look at this backward:

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

We can interpret this to mean that:

$$\frac{\Delta f}{\Delta x} \approx f'(x_0) \quad \text{when } \Delta x \approx 0.$$

In other words, the average rate of change $\frac{\Delta f}{\Delta x}$ is nearly the same as the infinitesimal rate of change $f'(x_0)$.

We can see that this is the same as our original formula

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

if we multiply both sides by Δx and remind ourselves what Δx and Δf are abbreviations for:

$$\begin{array}{rcl} \dfrac{\Delta f}{\Delta x} & \approx & f'(x_0) \\ \Delta x \cdot \dfrac{\Delta f}{\Delta x} & \approx & f'(x_0) \cdot \Delta x \\ \Delta f & \approx & f'(x_0) \cdot \Delta x \\ f(x) - f(x_0) & \approx & f'(x_0)(x - x_0) \\ f(x) & \approx & f(x_0) + f'(x_0)(x - x_0) \end{array}$$

So we have two different ways of writing a formula for linear approximation. When you're solving linear approximation problems, try to choose the most appropriate formula for the problem you're working on.

Approximations at 0 for Sine, Cosine and Exponential Functions

Here is a list of several linear approximations which you may want to memorize. Half the work of memorizing a linear approximation is memorizing the derivative of a function at a base point, so memorizing these formulas should improve your knowledge of derivatives.

To make things as simple as possible, we always use base point $x_0 = 0$ and assume that $x \approx 0$. Then our general formula becomes:

$$f(x) \approx f(0) + f'(0)x.$$

Remember that when x is not near zero, this approximation probably won't work.

(Later we'll discuss exactly how close x has to be to zero; this is partly a matter of intuition and is very important in applications.)

We want to find linear approximations for the functions $\sin x$, $\cos x$ and e^x when x is near 0. We'll start by building a table of values of $f'(x)$, $f(0)$, and $f'(0)$; from these we can "read off" the linear approximations.

$f(x)$	$f'(x)$	$f(0)$	$f'(0)$
$\sin x$	$\cos x$	0	1
$\cos x$	$-\sin x$	1	0
e^x	e^x	1	1

We can now plug the values for $f(0)$ and $f'(0)$ into our formula $f(x) \approx f(0) + f'(0)x$ to get linear approximations for these functions:

1. $\sin x \approx x \quad (\text{if } x \approx 0) \quad$ (see part (a) of Fig. 1)
2. $\cos x \approx 1 \quad (\text{if } x \approx 0) \quad$ (see part (b) of Fig. 1)
3. $e^x \approx 1 + x \quad (\text{if } x \approx 0)$

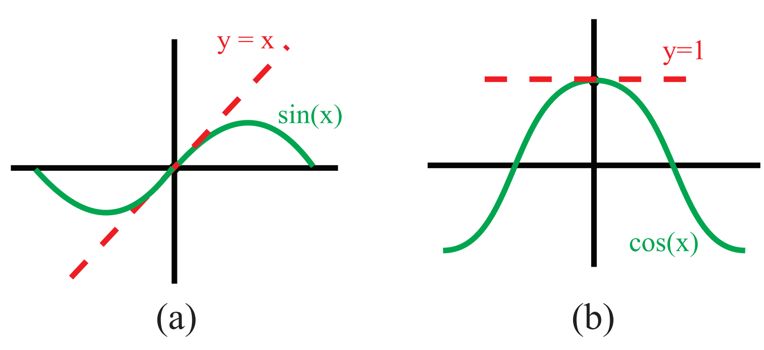


Figure 1: Linear approximations to sine and cosine at $x = 0$.

Approximations at 0 for $\ln(1+x)$ and $(1+x)^r$

Next, we compute two linear approximations that are slightly more challenging.

Here's the table of values:

$f(x)$	$f'(x)$	$f(0)$	$f'(0)$
$\ln(1+x)$	$\frac{1}{1+x}$	0	1
$(1+x)^r$	$r(1+x)^{r-1}$	1	r

And here are the linear approximations we get from the table:

1. $\ln(1+x) \approx x \quad (\text{if } x \approx 0)$
2. $(1+x)^r \approx 1 + rx \quad (\text{if } x \approx 0)$

Remember that we computed the linear approximation to $\ln x$ at $x_0 = 1$. Since our base point wasn't 0 we couldn't include that here. Because $\ln x \to -\infty$ as $x \to 0$, a linear approximation of $\ln x$ near $x_0 = 0$ is useless to us. Instead we have a linear approximation of the function $\ln(1+x)$ near our default base point $x_0 = 0$, which works out to nearly the same thing as a linear approximation of $\ln x$ near $x_0 = 1$.

Similarly, we found a linear approximation to $(1 + x)^r$; not to x^r. For some values of r, x^r is not well behaved when $x = 0$. If we really need an approximation of x^r we can get one by a change of variables.

For example, in a previous example we computed that $\ln u \approx u - 1$ for $u \approx 0$ (we've just replaced x by u.) Now we change variables by setting $u = 1 + x$. If we plug in $1 + x$ everywhere we had a u we get:

$$\ln(1+x) \approx (1+x) - 1 = x,$$

which is exactly the formula we have above.

If you've memorized $\ln(1+x) \approx x$ for $x \approx 0$ you can quickly find an approximation for $\ln u$ for $u \approx 1$ through the change of variables $x = u - 1$.

Curves are Hard, Lines are Easy

How are linear approximations used? We'll start with an example, then discuss.

Suppose I want to know the value of $\ln(1.1)$. If I've memorized the formula $\ln(1+x) = x$ for $x \approx 0$ I know right away that $\ln(1.1) \approx 0.1$ just by plugging in $x = 0.1$. (This works because 0.1 is "close enough" to zero.)

So what? The value of $\ln(1.1)$ is hard to compute; the value of 0.1 is easy. We used linear approximation to make a "hard" value "easy" to understand.

In general, $f(x)$ is hard to compute and $f(x_0) + f'(x_0)(x - x_0)$ is easy to compute (even though $f(x_0) + f'(x_0)(x - x_0)$ looks uglier). Look at the list below; evaluating the expressions on the left is hard and evaluating the ones on the right is easy.

$$\begin{aligned} \sin x &\approx x \\ \cos x &\approx 1 \\ e^x &\approx 1 + x \\ \ln(1+x) &\approx x \\ (1+x)^r &\approx 1 + rx \end{aligned}$$

That's the main advantage of linear approximation: it lets you work with expressions that are much easier to compute, which lets you make faster progress in solving problems.

Linear Approximation of a Complicated Exponential

Example 3: Find the linear approximation of $f(x) = \dfrac{e^{-3x}}{\sqrt{1+x}}$ near $x = 0$. We could calculate $f'(x)$ and find $f'(0)$. But instead, we will do this by algebraically combining the linear approximations we already have.

From our list of linear approximations we have:

$$\begin{array}{rcl} e^x & \approx & 1 + x \\ (1+x)^r & \approx & 1 + rx \end{array}$$

So $\frac{1}{\sqrt{(1+x)}} = (1+x)^{-1/2} \approx 1 - \frac{x}{2}$.

We rewrite $\dfrac{e^{-3x}}{\sqrt{1+x}}$ as $e^{-3x} \cdot (1+x)^{-1/2}$ so that all we'll have to do is multiply two linear equations to get our approximation.

$$e^{-3x} \approx 1 + (-3x)$$

$$\frac{1}{\sqrt{1+x}} = (1+x)^{-1/2} \approx 1 - \frac{1}{2}x$$

Put these two approximations together to get:

$$\begin{array}{rcl} \dfrac{e^{-3x}}{\sqrt{1+x}} & \approx & (1-3x)(1-\dfrac{1}{2}x) \\ & \approx & 1 - \dfrac{1}{2}x - 3x + \dfrac{3}{2}x^2 \\ & \approx & 1 - \dfrac{1}{2}x - \dfrac{6}{2}x + \dfrac{3}{2}x^2 \end{array}$$

Remember that x is a number close to 0, so x^2 is a number very close to 0. In fact, when working with linear approximations we assume that x^2 is so small that we can safely ignore it. Our approximation then becomes:

$$\frac{e^{-3x}}{\sqrt{1+x}} \quad \approx \quad 1 - \frac{7}{2}x$$

When you're allowed to ignore all the higher degree terms — x^2, x^3 and so on — this sort of algebra becomes very simple. Why are we allowed to do that here and not in our algebra classes? For one thing, we're not calculating exact values only approximations. We know that the graph of the tangent line probably isn't the same as the graph of the function; we've already decided that finding an exact value is too hard. For another thing, these approximations are only valid when x is close to 0. If x is $1/100$ then x^2 is $1/10000$. The $\frac{3}{2}x^2$ we discarded has a value of 0.00015 when $x = .01$. When x is "small enough", any terms involving x^2 are so small that they don't significantly affect the final estimate.

Question: Can we use the original formula?

Earlier, we found that:

$$f(x) = \frac{e^{-3x}}{\sqrt{1+x}} \approx 1 - \frac{7}{2}x.$$

Could we use a different method to get a linear approximation of the function $f(x)$?

Yes. We could calculate f' and use the formula for linear approximation to find:

$$f(x) \approx f(0) + f'(0)x.$$

This must also be a linear approximation to $\dfrac{e^{-3x}}{\sqrt{1+x}}$.

We can easily find that $f(0) = 1$. Computing $f'(x)$ by the product rule is an annoying, somewhat long computation. Because of what we've just done we know that $f'(0)$ must equal $-\frac{7}{2}$. We used linear approximation as a shortcut to avoid computing $f'(0)$ directly.

When we study quadratic approximation we'll quickly see that combining approximations for complicated functions is far superior to differentiating them twice.

Question: If we find the linear approximation by differentiating, do we have to throw away an x^2 term?

Answer: No. But remember that when x is close to 0 throwing away an x^2 term has very little influence on our final value. Throwing away the x^2 was an easy way to simplify our expression; it's not something we should be trying to avoid here. (Linear approximation just captures the linear features of the function; we are not concerning ourselves with higher order terms here.)

Applications of Linear Approximation

In this unit we're trying to learn about applications of the derivative to real problems. Here is one such example that involves math as well as physics.

Example 4: Planet Quirk

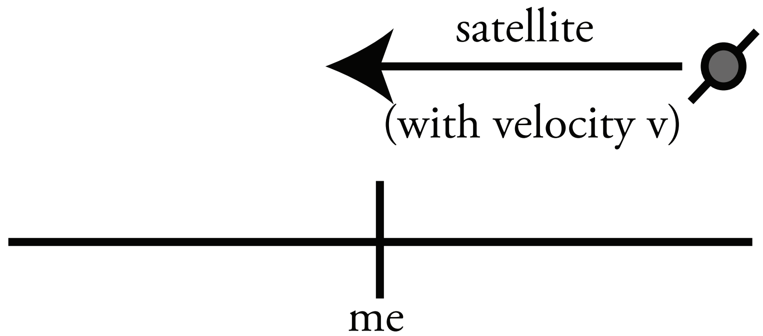


Figure 1: Illustration of Example 4: a satellite with velocity v speeding past "me" on planet Quirk.

Let's say I am on Planet Quirk, and that a satellite is whizzing overhead with a velocity v. The satellite has a clock on it that reports a time T. I have a watch that reports a time T_m. We want to calculate the time dilation (a concept from special relativity) that describes the difference between T and T_m.

We borrow the following equation from special relativity:

$$T_m = \frac{T}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Again, T_m is the time I measure on my wristwatch, and T is the time measured on board the satellite. How different are T_m and T?

To avoid dividing by a square root, we'll once again use linear approximation:

$$\frac{T}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{T}{\sqrt{1-u}} \approx T\left(1 + \frac{1}{2}u\right),$$

where $u = \frac{v^2}{c^2}$. (If you're wondering why we got $(1 + \frac{1}{2}u)$ and not $(1 - \frac{1}{2}u)$, try applying the change of variables $v = -u$ before approximating.)

And so, we find that:

$$T_m = \frac{T}{\sqrt{1 - \frac{v^2}{c^2}}} \approx T\left(1 + \frac{1}{2}\left(\frac{v^2}{c^2}\right)\right)$$

How does this affect you in the real world? GPS transmitters are mounted on satellites, the satellites are moving, and you might be moving too. According to special relativity, there will be a difference between the time on your GPS device and the time on the satellite. This time difference will affect the GPS device's estimate of your position.

The engineers who set up the GPS satellite system knew this and had to decide if they needed to take this into account when designing the system. It turns out that GPS satellites move at about $v = 4$ kilometers per second (km/s) and $c = 3 \times 10^5$ km/sec, and so $\frac{v^2}{c^2} \approx 10^{-10}$ and $T_m \approx T(1.00000000005)$. There's hardly any difference between the times measured on the ground and in the satellite.

Because $\frac{v^2}{c^2}$ is very close to zero, our linear approximation should be quite close to the actual value of T_m. Another good reason for using linear approximation here is that if the answer is "the difference is too small to matter", the person doing the calculation has no use for a more precise answer which may be more difficult to calculate.

Nevertheless, engineers used this very approximation (along with several other such approximations) to calibrate the radio transmitters on GPS satellites. The satellites transmit at a slightly offset frequency.

Relative Error

We continue with our example of time dilation in GPS satellite operation. We started with the following formula from special relativity:

$$T_m = \frac{T}{\sqrt{1 - \frac{v^2}{c^2}}}$$

and used a linear approximation to find that:

$$T_m \approx T\left(1 + \frac{1}{2}\left(\frac{v^2}{c^2}\right)\right).$$

This formula describes the difference due to time dilation between clocks on the ground and on the satellite. Algebraically, the difference is $\Delta T = T_m - T$; it turns out that there's a very simple relation between T, ΔT, v and c:

$$\begin{aligned}
T_m &\approx T\left(1 + \frac{1}{2}\left(\frac{v^2}{c^2}\right)\right) \\
T_m &\approx T + T\frac{1}{2}\left(\frac{v^2}{c^2}\right) \\
T_m - T &\approx T\frac{1}{2}\left(\frac{v^2}{c^2}\right) \\
\Delta T &\approx T\frac{1}{2}\left(\frac{v^2}{c^2}\right) \\
\frac{\Delta T}{T} &\approx \frac{1}{2}\left(\frac{v^2}{c^2}\right)
\end{aligned}$$

In other words, the relative or percent error $\frac{\Delta T}{T}$ caused by time dilation is proportional to the ratio $\dfrac{v^2}{c^2}$, which relates the speed of the satellite to the speed of light.

As in the example of falling stock prices, this value $\frac{\Delta T}{T}$ gives us an idea of the relative size of the error introduced by time dilation.

The Formula for Quadratic Approximation

Quadratic approximation is an extension of linear approximation – we're adding one more term, which is related to the second derivative. The formula for the quadratic approximation of a function $f(x)$ for values of x near x_0 is:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \quad (x \approx x_0)$$

Compare this to our old formula for the linear approximation of f:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \quad (x \approx x_0).$$

We got from the linear approximation to the quadratic one by adding one more term that is related to the second derivative:

$$f(x) \approx \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{Linear Part}} + \underbrace{\frac{f''(x_0)}{2}(x - x_0)^2}_{\text{Quadratic Part}} \quad (x \approx x_0)$$

These are more complicated and so are only used when higher accuracy is needed.

We'd like to develop a catalog of quadratic approximations similar to our catalog of linear approximations. Let's start by looking at the quadratic version of our estimate of $\ln(1.1)$. The formula for the quadratic approximation turns out to be:

$$\ln(1 + x) \approx x - \frac{x^2}{2},$$

and so $\ln(1.1) = \ln(1 + \frac{1}{10}) \approx \frac{1}{10} - \frac{1}{2}(\frac{1}{10})^2 = 0.095$. This is not the value 0.1 that we got from the linear approximation, but it's pretty close (and slightly more accurate).

Explaining the Formula by Example

As we saw last time, quadratic approximations are a little more complicated than linear approximation. Use these when the linear approximation is not enough. For example, most modeling in economics uses quadratic approximation. When using approximation you sacrifice some accuracy for the ability to perform complex calculations; using approximations more precise than quadratic approximations can make your calculations too unwieldy to be useful.

The basic formula for quadratic approximation with base point $x_0 = 0$ is:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \quad (x \approx x_0)$$

This works for values of x near 0; it is just the formula for linear approximation with one new term.

Where did that extra term come from, and what is the $\frac{1}{2}$ doing in front of the x^2? We'll explain this in terms of what happens if the graph of our function is a parabola. If the graph of a function is a parabola, that function *is* a quadratic function. Ideally, the quadratic approximation of a quadratic function should be identical to the original function.

For instance, consider:

$$f(x) = a + bx + cx^2; \qquad f'(x) = b + 2cx; \qquad f''(x) = 2c.$$

Set the base point $x_0 = 0$. Now we try to recover the values of a, b and c from our information about the derivatives of $f(x)$.

$$\begin{aligned} f(0) &= a + b \cdot 0 + c \cdot 0^2 &\implies& \quad a = f(0) \\ f'(0) &= b + 2c \cdot 0 = b &\implies& \quad b = f'(0) \\ f''(0) &= 2c &\implies& \quad c = \frac{f''(0)}{2} \end{aligned}$$

This tells us what the coefficients of the quadratic approximation formula must be in order for the quadratic approximation of a quadratic function to equal that function. To confirm this, we see that applying the formula:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \quad (x \approx x_0)$$

to our quadratic function $f(x) = a + bx + cx^2$ yields the quadratic approximation:

$$f(x) \approx a + bx + \frac{2c}{2}x^2.$$

Another way to think about this is that in the linear approximation of a function, the first derivative of the approximation is the same as the first derivative of the function. In a quadratic approximation, the first and second derivatives of the approximation are the same as the first and second derivatives of the function. If the $\frac{1}{2}$ weren't there, this wouldn't be true.

Quadratic Approximation at 0 for Several Examples

We'll save the derivation of the formula:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \quad (x \approx x_0)$$

for later; right now we're going to find formulas for quadratic approximations of the functions for which we have a library of linear approximations.

Basic Quadratic Approximations:

In order to find quadratic approximations we need to compute second derivatives of the functions we're interested in:

$f(x)$	$f'(x)$	$f''(x)$	$f(0)$	$f'(0)$	$f''(0)$
$\sin x$	$\cos x$	$-\sin x$	0	1	0
$\cos x$	$-\sin x$	$-\cos x$	1	0	-1
e^x	e^x	3^x	1	1	1
$\ln(1+x)$	$\frac{1}{1+x}$	$\frac{-1}{(1+x)^2}$	0	1	-1
$(1+x)^r$	$r(1+x)^{r-1}$	$r(r-1)(1+x)^{r-2}$	1	r	$r(r-1)$.

Plugging the values for $f(0)$, $f'(0)$ and $f''(0)$ in to the quadratic approximation we get:

1. $\sin x \approx x \quad (\text{if } x \approx 0)$

2. $\cos x \approx 1 - \frac{x^2}{2} \quad (\text{if } x \approx 0)$

3. $e^x \approx 1 + x + \frac{1}{2}x^2 \quad (\text{if } x \approx 0)$

4. $\ln(1+x) \approx x - \frac{1}{2}x^2 \quad (\text{if } x \approx 0)$

5. $(1+x)^r \approx 1 + rx + \frac{r(r-1)}{2}x^2 \quad (\text{if } x \approx 0)$

We've computed some formulas; now let's think about their meaning.

Geometric significance (of the quadratic term)

A quadratic approximation gives a best-fit parabola to a function. For example, let's consider $f(x) = \cos(x)$ (see Figure 1).

The linear approximation of $\cos x$ near $x_0 = 0$ approximates the graph of the cosine function by the straight horizontal line $y = 1$. This doesn't seem like a very good approximation.

The quadratic approximation to the graph of $\cos(x)$ is a parabola that opens downward; this is much closer to the shape of the graph at $x_0 = 0$ than the line

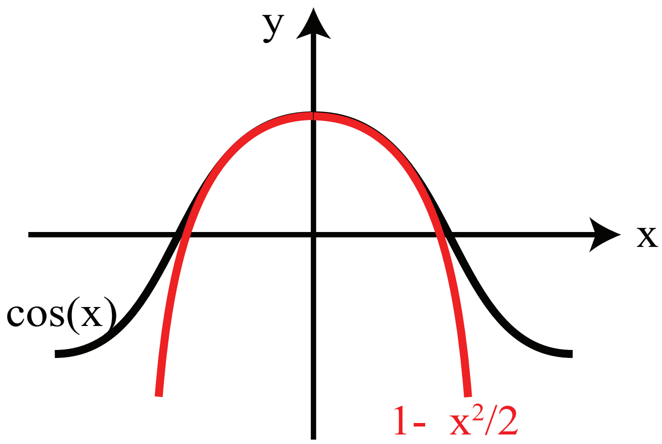


Figure 1: Quadratic approximation to $\cos(x)$.

$y = 1$. To find the equation of this quadratic approximation we set $x_0 = 0$ and perform the following calculations:

$$\begin{aligned} f(x) &= \cos(x) & \implies \quad & f(0) = \cos(0) = 1 \\ f'(x) &= -\sin(x) & \implies \quad & f'(0) = -\sin(0) = 0 \\ f''(x) &= -\cos(x) & \implies \quad & f''(0) = -\cos(0) = -1. \end{aligned}$$

We conclude that:

$$\cos(x) \approx 1 + 0 \cdot x - \frac{1}{2}x^2 = 1 - \frac{1}{2}x^2.$$

This is the closest (or "best fit") parabola to the graph of $\cos(x)$ when x is near 0.

Quadratic Approximation

Last class we derived a list of quadratic approximations for values of x near 0.

Using the formula:

$$f(x) \approx f(0) + f'(0)x + \frac{f''(0)}{2}x^2 \quad (x \approx 0)$$

We can (and eventually will) calculate the following approximations:

- $\sin x \approx x \quad (\text{if } x \approx 0)$
- $\cos x \approx 1 - \frac{x^2}{2} \quad (\text{if } x \approx 0)$
- $e^x \approx 1 + x + \frac{1}{2}x^2 \quad (\text{if } x \approx 0)$
- $\ln(1+x) \approx x - \frac{1}{2}x^2 \quad (\text{if } x \approx 0)$
- $(1+x)^r \approx 1 + rx + \frac{r(r-1)}{2}x^2 \quad (\text{if } x \approx 0)$

Eventually you will recognize and remember all of these formulas, but it may take time and practice.

We have not derived the final two approximations on this list; we'll use them in several examples then describe their derivation.

Approximation of $\ln e$

Here's an example of the power of linear approximation, and of what quadratic approximation can do for us that linear approximation cannot.

Recall that when we discussed exponential and logarithmic functions we said that:

$$a_k = \left(1 + \frac{1}{k}\right)^k$$

tends to e as k goes to infinity. We did that by taking the logarithm of both sides:

$$\ln a_k = k \ln\left(1 + \frac{1}{k}\right),$$

and then analyzing the limit on the right hand side. That was a fairly difficult calculation which is made much easier by linear approximation. Since the linear approximation of $\ln(1 + x)$ is just x,

$$\ln a_k = k \ln\left(1 + \frac{1}{k}\right) \approx k(1/k) = 1.$$

Can we conclude that $\ln a_k = 1$ so that $a_k = e$? (We used an approximation "$\approx$" in the above, not "$=$".) The linear approximation only works when x is near 0, but as k goes to infinity $\frac{1}{k}$ is indeed near 0. So as k approaches infinity, the linear approximation gets closer and closer to the exact value of the function, and $\ln a_k$ approaches 1. Linear approximation is often used in this way to evaluate limits.

Now if we want to find the *rate* of convergence — if we want to find out how fast the value of $k \ln\left(1 + \frac{1}{k}\right)$ approaches 1 — we need to look at the size of $\ln a_k - 1$ for large values of k. To do this you'll use quadratic approximation; the formula for the quadratic approximation of the natural log function was given above:

$$\ln(1 + x) \approx x - \frac{1}{2}x^2 \quad \text{(for } x \text{ near 0).}$$

You need the next higher order term to get a more detailed understanding of $\lim_{k\to\infty} \ln a_k$; this question will be on the problem set.

Question: How do we know when to use a linear approximation and when to use a quadratic one?

Answer: This is a very good question. For now the questions you get will specify whether to use a linear or a quadratic approximation. As time goes on you should try to get a feel for when you can get away with a linear approximation. In general, only use a quadratic approximation if a linear approximation won't work, because quadratic approximations are much more complicated (as you'll see in the next example).

In real life when you're faced with a problem like this — for example, determining the effects of gravity on an orbiting satellite — nobody is going to tell you anything. You won't even be told whether a linear approximation is relevant; you're on your own.

Example: $\dfrac{e^{-3x}}{\sqrt{1+x}}$

Last lecture we computed the linear approximation for x near 0 of

$$\frac{e^{-3x}}{\sqrt{1+x}} = e^{-3x}(1+x)^{-1/2}.$$

This lecture we'll compute a quadratic approximation for this function when x is near 0.

To do this we need to use the quadratic approximations for e^{-3x} and $(1+x)^{-1/2}$. We'll use the following two approximation formulas:

$$\begin{aligned} e^x &\approx 1 + x + \frac{1}{2}x^2 \\ (1+x)^r &\approx 1 + rx + \frac{r(r-1)}{2}x^2 \end{aligned}$$

substituting $x = 3x$ into the first and $r = -\frac{1}{2}$ into the second.

$$e^{-3x}(1+x)^{-1/2} \approx \left(1 + (-3x) + \frac{1}{2}(-3x)^2\right)\left(1 + \left(-\frac{1}{2}\right)x + \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{2}x^2\right)$$

This looks awful! But we can ignore any terms of higher degree than x^2 and avoid doing those multiplications when we apply the distributive law, so it's not as bad as it looks.

$$e^{-3x}(1+x)^{-1/2} \approx 1 - 3x - \frac{1}{2}x + \frac{3}{2}x^2 + \frac{9}{2}x^2 + \frac{3}{8}x^2$$

Now we combine like terms:

$$e^{-3x}(1+x)^{-1/2} \approx 1 - \frac{7}{2}x + \frac{51}{8}x^2$$

Remember that this approximation is only valid for $x \approx 0$, and notice that the first two terms are exactly the linear approximation we got last time.

As you can see, calculations with quadratic approximations are much more involved than those with just linear approximations.

Question: Why do we get to drop all the higher order terms?

Answer: Because in the situation in which we're going to apply this, x is a very small number like $\frac{1}{100}$. That means that $x^2 \approx \frac{1}{10000}$ and $x^3 \approx \frac{1}{1000000}$. We don't need an exact answer so we can safely ignore anything as small as a millionth, which is what our x^3 terms represent.

Approximating $\ln(1+x)$ and $(1+x)^r$

- $\sin x \approx x \quad$ (if $x \approx 0$)
- $\cos x \approx 1 - \dfrac{x^2}{2} \quad$ (if $x \approx 0$)
- $e^x \approx 1 + x + \dfrac{1}{2}x^2 \quad$ (if $x \approx 0$)
- $\ln(1+x) \approx x - \dfrac{1}{2}x^2 \quad$ (if $x \approx 0$)
- $(1+x)^r \approx 1 + rx + \dfrac{r(r-1)}{2}x^2 \quad$ (if $x \approx 0$)

Now that we've seen a couple of examples of quadratic approximation, we'll derive the last two formulas in our library, shown above. The general formula for a quadratic approximation is:

$$f(x) \approx f(0) + f'(0)x + \frac{f''(0)}{2}x^2 \quad (x \approx 0)$$

As usual, we chose the base point $x_0 = 0$. Shown below are the first and second derivatives of the functions we're interested in and their values at $x_0 = 0$. Combining this with the general formula yields the quadratic approximations listed above.

$f(x)$	$f'(x)$	$f''(x)$	$f(0)$	$f'(0)$	$f''(0)$
$\sin x$	$\cos x$	$-\sin x$	0	1	0
$\cos x$	$-\sin x$	$-\cos x$	1	0	-1
e^x	e^x	3^x	1	1	1
$\ln(1+x)$	$\frac{1}{1+x}$	$\frac{-1}{(1+x)^2}$	0	1	-1
$(1+x)^r$	$r(1+x)^{r-1}$	$r(r-1)(1+x)^{r-2}$	1	r	$r(r-1)$

We can approximate most common functions using algebraic combinations of the functions in this library.

Introduction to Curve Sketching

Goal: To draw the graph of f using information about whether f' and f'' are positive or negative. We want the graph to be qualitatively correct, but not necessarily to scale.

WARNING: Don't abandon your precalculus skills and common sense — you already know a lot about graphing functions; calculus just fills in the gaps.

The first principle is that if f' is positive, then f is increasing. In other words, if the tangent line is pointing up then the function is going up too. Similarly, if the derivative of f is negative then f is decreasing.

The second principle is just a second order effect of the same type. If f'' is positive, that means that f' is increasing. This is just the first principle applied to the second derivative; f'' is the derivative of f'.

Figure 1 shows the graph of a function for which the second derivative is positive. You can see from the tangent lines sketched on the graph that their slopes increase from negative on the left to positive on the right. We say that curves with this shape are *concave up*.

Similarly, if $f'' < 0$ then f' is decreasing and the graph of f is *concave down*.

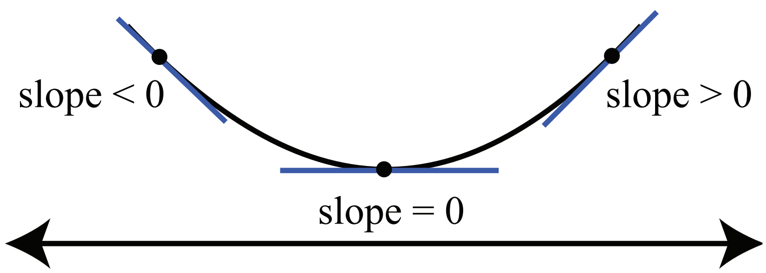


Figure 1: f is concave-up. The slope increases from negative to positive as x increases.

Curve Sketching Example 1

Example 1: Sketch the graph of $f(x) = 3x - x^3$.

First we note that:

$$f'(x) = 3 - 3x^2 = 3(1 - x)(1 + x)$$

We can see that if $-1 < x < 1$, both $(1 - x)$ and $(1 + x)$ are positive, so $f'(x)$ must also be positive, so f is increasing (by our first principle.)

When $x > 1$, $f'(x) < 0$ and f is decreasing. When $x < -1$, then $(1 - x)$ is positive and $(1 + x)$ is negative, so $f'(x) = 3(1 - x)(1 + x)$ is negative and f is decreasing.

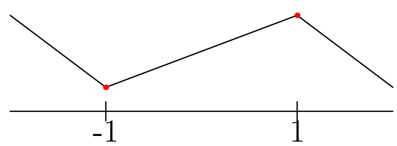


Figure 1: Turning points of $f(x) = 3x - x^3$.

We get a rough schematic of the graph of the function by drawing a number line at the bottom of our page as shown in Figure 1. Above the interval from negative infinity to -1, draw a diagonal line slanting down and to the right — this is one of the intervals on which f is decreasing. Above the interval from -1 to 1 draw a line slanting up and to the right — f is increasing. Finally, draw a line slanting down and to the right above the interval from 1 to positive infinity. The resulting zigzag gives us an idea of what an accurately drawn graph might look like.

We immediately notice two important features of the function lie at the points where the graph changes direction. These places where the derivative changes sign are turning points in the graph.

Definition: If $f'(x_0) = 0$, we call x_0 a *critical point* and $y_0 = f(x_0)$ is a *critical value* of f.

Our next step in understanding the graph of $f(x)$ is to plot the critical points and values of f. The critical points are $x = 1$ and $x = -1$; these are the values of x for which $(x - 1)$ and $(x + 1)$ are zero. The critical values are $f(1) = 3 \cdot 1 - 1^3 = 2$ and $f(-1) = 3 \cdot (-1) - (-1)^3 = -2$.

We plot the two points $(-1, -2)$ and $(1, 2)$, which we know are on the graph of f. Because we know where f' is positive and where it is negative, we also know that the graph decreases to $(-1, -2)$ and then starts to increase and that it increases toward $(1, 2)$ and then starts to decrease. In other words, the graph is shaped like a smile near $(-1, -2)$ and like a frown near $(1, 2)$. (We could also learn this from the second derivative.)

If we notice that $f(0) = 0$, we can now guess what the entire graph might look like but we still don't know for sure how far it decreases or increases to the left and right.

We can also notice that because all the powers of x are odd, the function f is odd; $f(-x) = 3(-x) - (-x)^3 = -3x + x^3 = -f(x)$. This means that if we can graph the function accurately for $x > 0$ we can reflect the graph across the y-axis to get a graph of the entire function.

In general, you should use precalculus skills to get information like $f(0) = 0$ or "f is odd" as much as you can.

The final detail we need to worry about is what happens at the "ends" of the graph. This topic is often neglected, but if you're using a graphing calculator or computer this can be the hardest part of understanding the graph of a function.

We want to know what happens as $x \to \infty$ and as $x \to -\infty$. As $x \to \infty$, the value of $-x^3$ grows very rapidly while the value of $3x$ is much smaller. So:

$$f(x) \approx -x^3 \to -\infty \text{ as } x \to \infty.$$

Similarly,

$$f(x) \approx x^3 \to \infty \text{ as } x \to -\infty.$$

We now know how far $f(x)$ increases or decreases as $x \to \pm\infty$ — it goes off toward infinity rather than, say, hugging the x-axis.

We now know a lot about the shape of the graph, but we can learn a little bit more about it by looking at the second derivative $f''(x) = -6x$. From this we learn that $f''(x) < 0$ when $x > 0$ and $f''(x) > 0$ when $x < 0$, so the graph is concave down to the right of the y-axis and concave up on the left. The value $x = 0$ is of interest not only because $f(0) = 0$ but also because it is an *inflection point* — a value x_0 for which $f''(x_0) = 0$.

We can now combine everything we've learned to get something like the graph shown in Figure 2.

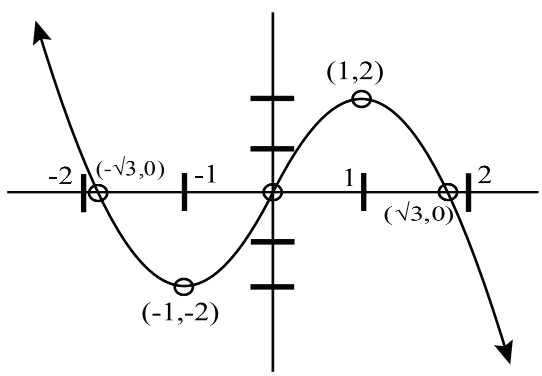


Figure 2: Sketch of the function $y = 3x - x^3$.

Question: What if the graph had a sharp point like the one in the schematic?

Answer: Points like that aren't called critical points, but they are very important. We'll talk about them later.

Graph Sketching Example

Example: Sketch the graph of $f(x) = \frac{x+1}{x+2}$.

What does the derivative of this function tell us about its graph? To save time, I'll tell you that

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

The special thing about this example is that $f'(x) \neq 0$, i.e. there are no values of x_0 for which $f(x_0) = 0$ and there are no critical points. If you've been trained that the first step in sketching graphs is to find out where the derivative is zero, you might be tempted to give up at this point. But we're not going to give up — can you think of anything you'd like to try here?

Student: Find values of x where f is undefined.

That's a sophisticated way of putting the point that I want to make, which is that we should go back to our precalculus skills and just plot points.

It turns out that the most important point to plot is the one that's not there: the one with x coordinate -2. This is what was just suggested — this is a point where the function is not defined.

If we try to compute $f(-2)$ we find that we can't because that makes the denominator 0. Instead we'll look at the limit as x approaches -2 from the left and from the right. We'll write a^+ for "a number x that is a little big bigger than a" to make these limits easier to talk about.

$$\lim_{x \to -2^+} f(x) = f(-2^+) = \frac{-2+1}{(-2)^+ + 2} = -10^+ = -\infty.$$

We see that the limit comes down to dividing -1 by a very small positive number; the result of this division is a very large negative number. From the other direction we get:

$$\lim_{x \to -2^-} f(x) = f(-2^-) = \frac{-2+1}{(-2)^- + 2} = -10^- = +\infty$$

This tells us what the graph of the function is doing near $x = 2$. We could evaluate the function at any point; this was the most interesting point.

Our next step is to consider the "ends" of the graph, where $x \to +\infty$ and $x \to -\infty$. This tells us what happens to the left and right of the "screen" of our graph, just as the previous calculation told us what was happening above and below.

A simple algebraic trick makes it easy to see what will happen to the value of $f(x)$ as $x \to \pm\infty$:

$$\begin{aligned} f(x) &= \frac{x+1}{x+2} \\ &= \frac{x+1}{x+2}\frac{1/x}{1/x} \\ f(x) &= \frac{1+\frac{1}{x}}{1+\frac{2}{x}} \end{aligned}$$

$$\lim_{x\to\pm\infty} f(x) \quad = \quad \frac{1}{1} = 1$$

As x gets huge, the values $\frac{1}{x}$ and $\frac{2}{x}$ approach 0. We could abbreviate this conclusion as $f(\pm\infty) = 1$.

To draw the graph we use asymptotes; there's a horizontal asymptote at $y = 1$ and a vertical one at $x = -2$. We use dotted lines to sketch these asymptotes so that we won't confuse them with the sketch of our graph as in Figure 1.

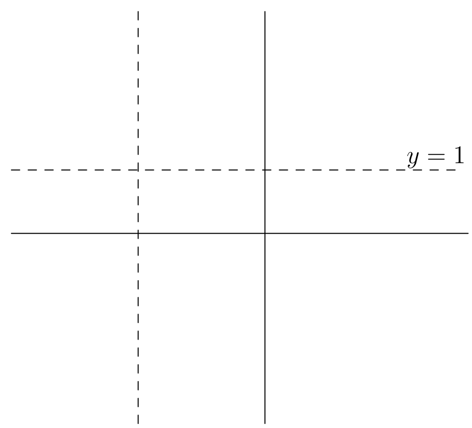


Figure 1: Axes and asymptotes.

The fact that $\lim_{x\to -2^+} f(x) = -\infty$ tells us that the graph plunges down as it approaches $x = -2$ from the right. As x approaches -2 from the left, the graph goes up toward positive infinity on the other side of the vertical asymptote. As x approaches negative infinity, the graph gets closer and closer to the horizontal asymptote $y = 1$, as it does when $x \to +\infty$.

We can almost finish this graph now, but there's one outstanding question. How do we know whether the graph crosses the line $y = 1$ or not?

Student: It can't dip below the line because there are no critical points.

That's exactly right. Because f' is never 0 the graph of f can't double back on itself, because the graph can't have any horizontal tangent lines.

We've practically got the entire graph now. We could improve it by adding details like the exact x- and y-intercepts, but we already understand the qualitative behavior of this function.

Let's look at the sign of the first derivative of the function to double check our graphing. To compute the derivative we can use the following algebraic trick:

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

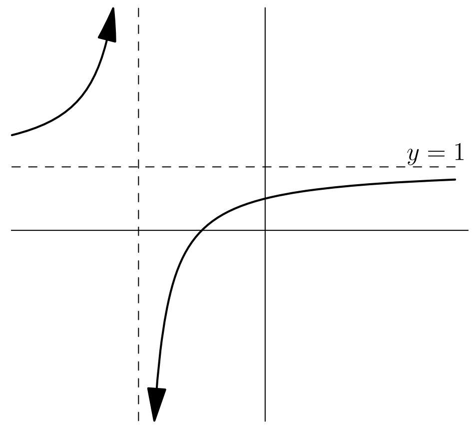


Figure 2: Sketch of $y = \dfrac{x+1}{x+2}$.

So $f'(x) = 0 - (x+2)^{-2} = \frac{1}{(x+2)^2}$ for $x \neq -2$. (This calculation also shows us that the graph of $f(x)$ is a hyperbola similar to the graph of $\frac{1}{x}$.)

We can see from our graph that $f(x)$ is increasing from negative infinity to -2 and from -2 to positive infinity, so $f'(x)$ must be positive. The derivative we calculated is always positive, so we've successfully checked that result.

Despite the fact that we know that f is increasing on $(-\infty, 2)$ and $(2, \infty)$, we shouldn't think that f is increasing for all x. There's a break at $x = -2$, which is the most important feature of the function.

The second derivative of $f(x)$ is

$$f''(x) = \frac{-2}{(x+2)^3} \quad (x \neq -2).$$

This is negative when $(x+2)^3$ is positive, so $f''(x) < 0$ when $x > -2$ and the graph is concave down on the interval $(-2, \infty)$. The second derivative is positive when $(x+2)^3$ is negative, so the graph is concave up on the interval $(-\infty, -2)$. Again, this agrees with the graph we have drawn. It also confirms that the graph doesn't "wiggle".

Question: Are we defining "increasing" to mean that the first derivative is positive?

Answer: No. Whenever f' is positive the function is increasing, but it's possible for the function to be increasing even when f' is 0.

General Strategy for Curve Sketching

1. (Precalc skill) Plot

   a discontinuities of f (especially infinite ones)

   b endpoints (or $x \to \pm\infty$)

   c easy points (optional)

2. Find the critical points — usually where the slope changes from positive to negative, or vice versa.

   a Solve $f'(x) = 0$

   b Plot critical points and values, but only if it's relatively easy to do so.

3. Decide whether $f'(x) < 0$ or $f'(x) > 0$ on each interval between critical points and discontinuities. (This just double checks steps 1 and 2.)

4. Decide whether $f''(x) < 0$ or $f''(x) > 0$ on each interval between critical points and discontinuities. This tells us whether the graph is concave up or concave down. Inflection points occur when $f''(x_0) = 0$. (If you can, skip this step.)

5. Combine this information to draw the graph.

Detailed Example of Curve Sketching

Example Sketch the graph of $f(x) = \dfrac{x}{\ln x}$. (Note: this function is only defined for $x > 0$)

1. Plot

   a The function is discontinuous at $x = 1$, because $\ln 1 = 0$.

   $$f(1^+) = \frac{1}{\ln 1^+} = \frac{1}{0^+} = \infty$$

   $$f(1^-) = \frac{1}{\ln 1^-} = \frac{1}{0^-} = -\infty$$

   b endpoints (or $x \to \pm\infty$)

   $$f(0^+) = \frac{0^+}{\ln 0^+} = \frac{0^+}{-\infty} = 0$$

   The situation is a little more complicated at the other end; we'll get a feel for what happens by plugging in $x = 10^{10}$.

   $$f(10^{10}) = \frac{10^{10}}{\ln 10^{10}} = \frac{10^{10}}{10 \ln 10} = \frac{10^9}{\ln 10} >> 1$$

   We conclude that $f(\infty) = \infty$.

We can now start sketching our graph. The point $(0, 0)$ is one endpoint of the graph. There's a vertical asymptote at $x = 1$, and the graph is descending before and after the asymptote. Finally, we know that $f(x)$ increases to positive infinity as x does. We already have a pretty good idea of what to expect from this graph!

2. Find the critical points

$$\begin{aligned} f'(x) &= \frac{1 \cdot \ln x - x\left(\frac{1}{x}\right)}{(\ln x)^2} \\ &= \frac{(\ln x) - 1}{(\ln x)^2} \end{aligned}$$

   a $f'(x) = 0$ when $\ln x = 1$, so when $x = e$. This is our only critical point.

   b $f(e) = \frac{e}{\ln e} = e$ is our critical value. The point (e, e) is a critical point on our graph; we can label it with the letter c. (It's ok if our graph is not to scale; we'll do the best we can.)

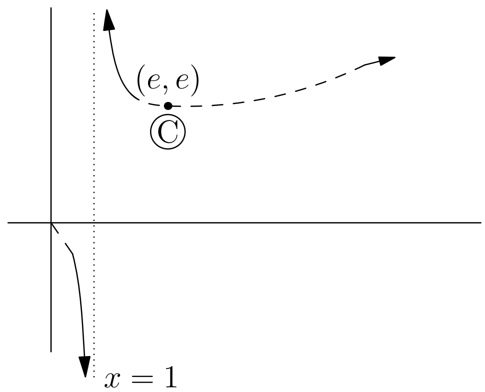


Figure 1: Sketch using starting point, asymptote, critical point and endpoints.

We now know the qualitative behavior of the graph. We know exactly where f is increasing and decreasing because the graph can only change direction at critical points and discontinuities; we've identified all of those. The rest is more or less decoration.

3. Double check using the sign of f'.

   We already know:

$$\begin{array}{ccc} f & \text{is decreasing on} & 0 < x < 1 \\ f & \text{is decreasing on} & 1 < x < e \\ f & \text{is increasing on} & e < x < \infty \end{array}$$

   We now double check this.

$$f'(x) = \frac{(\ln x) - 1}{(\ln x)^2}$$

   When x is between 0 and 1, $f'(x)$ equals a negative number divided by a positive number so is negative.

   When x is between 1 and e, $f'(x)$ again equals a negative number divided by a positive number so is negative.

   When x is between e and ∞, $f'(x)$ equals a positive number divided by a positive number so is positive.

   This confirms what we learned in steps 1 and 2.

   Sometimes steps 1 and 2 will be harder; then you might need to do this step first to get a feel for what the graph looks like.

   There's one more piece of information we can get from the first derivative $f'(x) = \frac{(\ln x)-1}{(\ln x)^2}$. It's possible for the denominator to be infinite; this is another situation in which the derivative is zero. So $f'(0^+) = 0$ and $x = 0^+$ is another critical point with critical value $-\infty$.

An easier way to see this is to rewrite $f'(x)$ as:

$$\frac{1}{\ln x} - \frac{1}{(\ln x)^2}$$

and note that:

$$f'(0^+) = \frac{1}{\ln 0^+} - \frac{1}{(\ln 0^+)^2} = \frac{1}{-\infty} - \frac{1}{(\infty)^2} = 0 - 0 = 0.$$

4. Use $f''(x)$ to find out whether the graph is concave up or concave down.

$$f'(x) = (\ln x)^{-1} - (\ln x)^{-2}$$

So

$$\begin{aligned} f''(x) &= -(\ln x)^{-2}\frac{1}{x} - -2(\ln x)^{-3}\frac{1}{x} \\ &= \frac{-(\ln x)^{-2} + 2(\ln x)^{-3}}{x}\frac{(\ln x)^3}{(\ln x)^3} \\ &= \frac{-\ln x + 2}{x(\ln x)^3} \\ f''(x) &= \frac{2 - \ln x}{x(\ln x)^3} \end{aligned}$$

We need to figure out where this is positive or negative. There are two places where the sign might change – when $2 - \ln x$ changes sign or when $(\ln x)^3$ changes sign. (Remember x will always be positive.)

The value of $2 - \ln x$ is positive when $\ln x < 2$ (when $x < e^2$) and negative when $x > e^2$. The denominator is positive when $x > 1$ and negative when $x < 1$. Combining these, we get:

$$\begin{aligned} 0 < x < 1 &\implies f''(x) < 0 \text{ (concave down)} \\ 1 < x < e^2 &\implies f''(x) > 0 \text{ (concave up)} \\ e^2 < x < \infty &\implies f''(x) < 0 \text{ (concave down)} \end{aligned}$$

This means that there's a "wiggle" at the point $\left(e^2, \frac{e^2}{2}\right)$ on the graph. The value of $f(x)$ is still increasing and the graph continues to rise, but the graph is rising less and less steeply as the values of $f'(x)$ decrease.

5. Combine this information to draw the graph.

   We've been doing this as we go. If you're working a homework problem, at this point you might copy your graph to a clean sheet of paper.

This is probably as detailed a graph as we'll ever draw. In fact, one advantage of our next topic is that it will reduce the need to be this detailed.

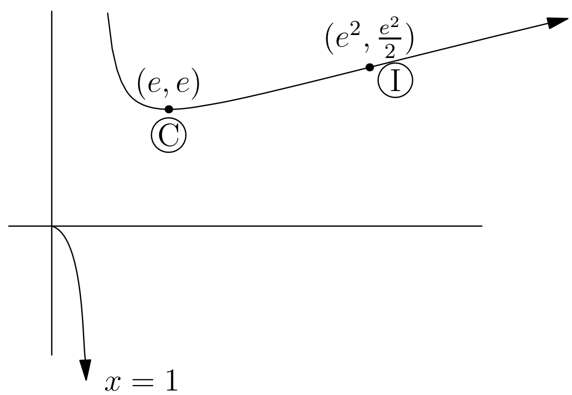


Figure 2: Final sketch.

Introduction to Maxima and Minima

Suppose you have a function like the one in Figure 1. Find the maximum value of the function. Then find the minimum.

To find the maximum value the function could output, we look at the graph and find the highest point. To find the lowest possible value we find the lowest point on the graph.

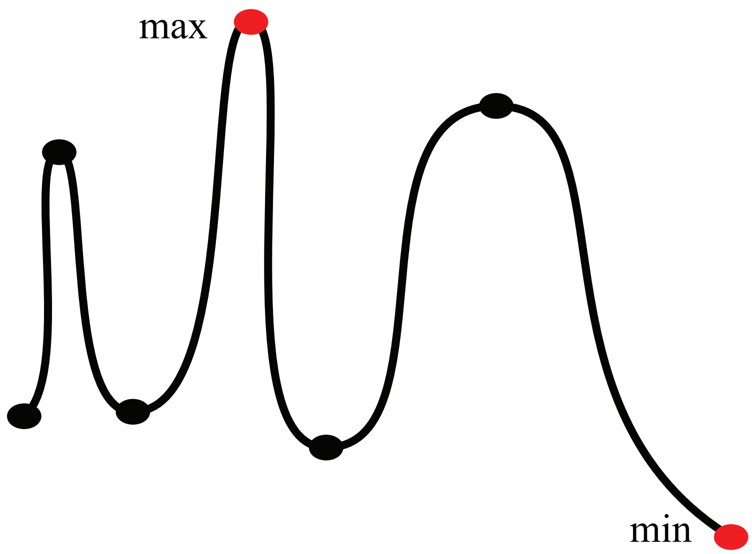


Figure 1: Search for max and min among critical points and endpoints

If you have a sketch, it is very easy to find max and the min. The problem is that the sketch is a lot of work. We don't want to do all that work every single time we need to find a maximum and minimum. Our goal is to use shortcuts; all we need to know is whether the graph is up or down.

Key to Finding Maxima and Minima: We only need to look at critical points *and* endpoints *and* points of discontinuity.

Looking back at our graph of $f(x) = \frac{x}{\ln x}$ we see five places to look for maxima and minima: $x = 0^+$, $x = 1^-$, $x = 1^+$, $x = e$ and $x = \infty$. Of these five points, only one is a critical point; we can't allow ourselves to forget about endpoints and points of discontinuity.

Maximum Area of Two Squares

Consider a wire of length 1, cut into two pieces. Bend each piece into a square. We want to figure out where to cut the wire in order to enclose as much area in the two squares as possible.

In all of these problems you start with a "bunch of words" — a story problem. The two main tasks in starting the problem are to draw a diagram and to pick variables.

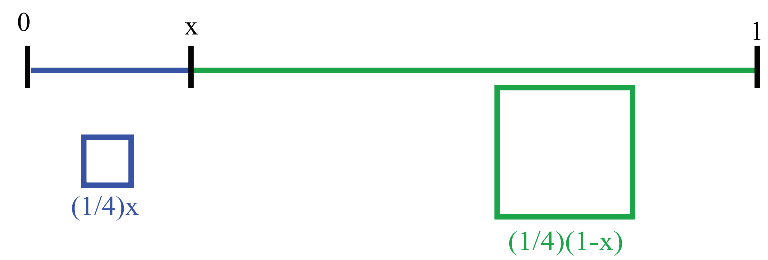


Figure 1: Two pieces of wire enclose two squares.

If we cut the wire so that one piece has length x, the other piece will have length $1-x$. We know that we're bending these pieces of wire into squares, so we can add those squares to our diagram. The first square will have sides of length $\frac{x}{4}$ and the second square will have sides of length $\frac{1-x}{4}$.

We want to find a maximum area, so we'll need formulas for the areas of these squares. The first square's area is $\frac{x^2}{16}$ and the second square has area $\left(\frac{1-x}{4}\right)^2$. The total area is then

$$\begin{aligned} A &= \left(\frac{x}{4}\right)^2 + \left(\frac{1-x}{4}\right)^2 \\ &= \frac{x^2}{16} + \frac{(1-x)^2}{16} \end{aligned}$$

Most calculus students' instinct at this point is to find the critical points — the values x_0 for which $A'(x_0) = 0$. We can do that now if you like:

$$\begin{aligned} A' &= \frac{2x}{16} + \frac{2(1-x)}{16}(-1) \\ &= \frac{x}{8} - \frac{1}{8} + \frac{x}{8} \\ A' &= \frac{2x-1}{8} \\ A' &= 0 \implies 2x - 1 = 0 \implies x = \frac{1}{2} \end{aligned}$$

So $x = \frac{1}{2}$ is a critical point with critical value:

$$A\left(\frac{1}{2}\right) = \left(\frac{\frac{1}{2}}{4}\right)^2 + \left(\frac{\frac{1}{2}}{4}\right)^2 = \frac{1}{32}$$

We're not done yet, though. We still need to check the endpoints! The length x of the first piece of wire has to satisfy $0 < x < 1$, so we should check the limits as x approaches the endpoints. In this case we can find those limits just by plugging in values. At $x = 0$,

$$A(0^+) = 0^2 + \left(\frac{1-0}{4}\right)^2 = \frac{1}{16}.$$

At $x = 1$,

$$A(1^-) = \left(\frac{1}{4}\right)^2 + 0^2 = \frac{1}{16}.$$

If we try to graph the function with the information we have now, we see that it starts at the point $\left(0, \frac{1}{16}\right)$, dips down to the point $\left(\frac{1}{2}, \frac{1}{32}\right)$, then goes back up to $\left(1, \frac{1}{16}\right)$. (See Fig. **??**.)

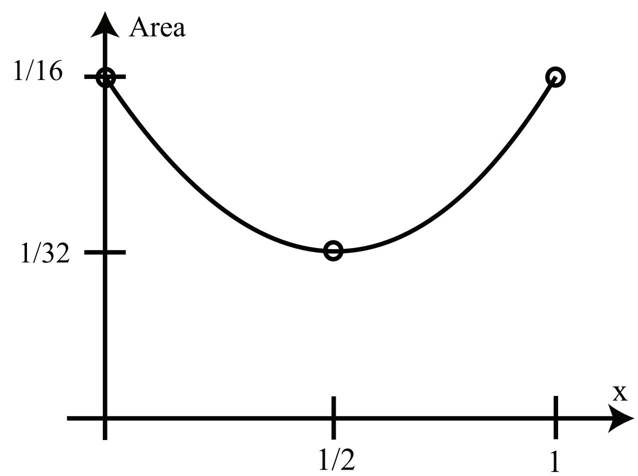


Figure 2: Graph of the area function.

When we found the critical point we did not find the maximum enclosed area — the *minimum* area was achieved at $x = \frac{1}{2}$. The maximum area is not achieved in $0 < x < 1$, but it is achieved at $x = 0$ or 1. The maximum corresponds to using the whole length of wire for one square.

Moral: If you don't pay attention to what the function looks like you may find the worst answer, rather than the best one.

We conclude that the least area enclosed by the two squares is $\frac{1}{32}$, when $x = \frac{1}{2}$; i.e when the two squares are equal. The greatest area enclosed is $\frac{1}{16}$ when $x = 0$ or $x = 1$ and there is only one square.

Commonly asked questions:

What is the minimum? It's the minimum value $\frac{1}{32}$.

Where is the minimum? It's at the critical point $x = \frac{1}{2}$.

These are two different questions. Be sure to answer the correct one; you may get so involved in doing the calculus to find $\frac{1}{2}$ that you forget to find the minimum value $\frac{1}{32}$. Both the critical point and the critical value are important; together they form the point on the graph $(\frac{1}{2}, \frac{1}{32})$ where it turns around.

There are many more and less precise ways to ask these questions. You'll have to do your best to understand what the questions (and answers) mean from the context of the question or example.

Question: Since the goal was to enclose as much area as possible, why did we find the minimum area?

Answer: The reason is that when we go about our procedure of looking for the least or the most, we'll automatically find both. We won't know which one is which until we compare values. It's actually to your advantage to figure out both the maximum and minimum whenever you answer such a question; otherwise you won't understand the behavior of the function very well.

Question: Could we also use the second derivative test here?

Answer: Yes, and we're going to see an example of the second derivative test soon. We could also look at the equation of $A(x)$ and notice that the graph must be a parabola that opens upward.

Max/Min Example 2

This is an example of a minimization problem with a constraint.

Example: Find the box (without a top) with least surface area for a fixed volume.

Again, we start by drawing a diagram and choosing variables. (This time we'll have four variable names.) I'll make things simpler by telling you in advance that the best box has a square bottom — knowing that the length of the box equals its width will let us use one fewer variable. It's often true that the correct answer is something symmetric, but you weren't expected to know this in advance.

Draw a picture of an open-topped box whose length and width equals x and whose height is y. The volume of the box is $V = x^2y$ and its surface area is the area of the base plus the area of each of the four sides: $A = x^2 + 4(xy)$.

The big difference between this problem and the last is that we have the *constraint* that the box must have a certain volume — this determines the relationship between x and y.

$$y = \frac{V}{x^2}$$

We can use this to rewrite the formula for A in terms of a single variable:

$$\begin{aligned} A(x) &= x^2 + 4x\frac{V}{x^2} \\ A(x) &= x^2 + \frac{4V}{x} \end{aligned}$$

Now that we have a formula for the value we're trying to minimize — the surface area — we'll follow the same procedure as before. Namely, we'll look for the critical points then check the endpoints and any discontinuities.

To find the critical points we take the derivative of $A(x)$ and set it equal to zero.

$$\begin{aligned} A(x) &= x^2 + \frac{4V}{x} \\ A'(x) &= 2x - \frac{4V}{x^2} \end{aligned}$$

$$\begin{aligned} 2x - \frac{4V}{x^2} &= 0 \\ 2x &= \frac{4V}{x^2} \\ x &= \frac{2V}{x^2} \\ x^3 &= 2V \\ x &= 2^{\frac{1}{3}}V^{\frac{1}{3}} \quad \text{(critical point)} \end{aligned}$$

We are not done; we don't even know whether this critical point gives us the most surface area or the least.

Let's check the ends: what are they? What's the smallest possible value of x?

Student: $x > 0$

This is a good answer! We can't build a box with a negative side length.

What's the largest possible value of x? It's true that:

$$x < \sqrt{\frac{V}{y}}$$

but since we said $y = \frac{V}{x^2}$ we can't use this as a limit.

Student: x is less than infinity?

That's right. $0 < x < \infty$.

This is an important realization; if there's no obvious limit on a variable the upper limit is infinity. And infinity is a very important endpoint, and is usually an easy endpoint to check.

We have to consider the possibility that if we shrink the sides all the way to $x = 0$ we'll get a better box. It would be very strange — maybe infinitely tall — but it might have the least surface area; we'll have to see.

$$A(0^+) = \left(x^2 + \frac{4V}{x}\right)\bigg|_{x=0^+}$$

Here the term $\frac{4V}{x}$ is going to infinity as x approaches 0 from the positive side; a box whose surface area approaches infinity is a bad box.

At the other end we get:

$$A(\infty) = \left(x^2 + \frac{4V}{x}\right)\bigg|_{x=\infty}$$

which is also infinite.

We can now draw a schematic for our graph. The surface area goes to infinity as x goes to 0 (from the positive side) and as x goes to infinity. There's only one critical point — only one place where the graph can turn around — so that point must be where the surface area reaches its minimum.

An alternative to checking ends is the *second derivative test*. The second derivative test is not recommended, and for most of the problems in this class it will be hard to use the second derivative test. However, this is a simple enough problem that we are able to use it.

$$\begin{aligned} A'(x) &= 2x - \frac{4V}{x^2} \\ A''(x) &= 2 + \frac{8V}{x^3} \end{aligned}$$

The second derivative is always positive (because x is always positive), so the graph of $A(x)$ is always concave up. This tells us that the critical point is at the lowest point of a "smile shape" and so must correspond to a minimum value of $A(x)$.

Question: Is this the answer to the question, or do we have to calculate y and A and so on?

Answer: What answer is appropriate depends on what the question is. We know that $x = 2^{\frac{1}{3}} V^{\frac{1}{3}}$. If we're going to build to build the box we'll also need to know the y value, which is the height of the box.

$$\begin{aligned} y &= \frac{V}{x^2} \\ &= \frac{V}{(2^{\frac{1}{3}} V^{\frac{1}{3}})^2} \\ y &= 2^{-\frac{2}{3}} V^{\frac{1}{3}} \end{aligned}$$

We could figure out the value of A; this would be appropriate if we wanted to know how much money it was going to cost to build this box.

$$\begin{aligned} A(x) &= x^2 + \frac{4V}{x} \\ &= (2^{\frac{1}{3}} V^{\frac{1}{3}})^2 + \frac{4V}{2^{\frac{1}{3}} V^{\frac{1}{3}}} \\ &= 2^{\frac{2}{3}} V^{\frac{2}{3}} + 4V \cdot 2^{-\frac{1}{3}} V^{-\frac{1}{3}} \\ &= 2^{\frac{2}{3}} V^{\frac{2}{3}} + (2 \cdot 2) 2^{-\frac{1}{3}} V^{\frac{2}{3}} \\ &= (2^{\frac{2}{3}} + 2 \cdot 2^{\frac{2}{3}}) V^{\frac{2}{3}} \\ &= 3 \cdot 2^{\frac{2}{3}} V^{\frac{2}{3}} \end{aligned}$$

So one possible answer to this problem is that the box with minimum surface area has:

$$\begin{aligned} \text{length} &= 2^{\frac{1}{3}} V^{\frac{1}{3}} \\ \text{height} &= 2^{-\frac{2}{3}} V^{\frac{1}{3}} \text{ and} \\ \text{surface area} &= 3 \cdot 2^{\frac{2}{3}} V^{\frac{2}{3}} \end{aligned}$$

However, there are much more meaningful ways of answering this question. If we use *dimensionless variables* it won't make a difference whether the sides of the box are measured in inches or kilometers and we may learn more about the problem. One famous dimensionless quantity is π, the ratio of the circumference of a circle to its diameter.

For example, $\frac{A}{V^{\frac{2}{3}}} = 3{\cdot}2^{\frac{1}{3}}$ is a dimensionless quantity, because if A is measured in $\in^2$ and V is measured in $\in^3$ the dimensions will cancel. If you increase the volume, you'll increase the surface area by the $\frac{2}{3}$ power of the volume.

The other dimensionless quantity is:

$$\frac{x}{y} = \frac{2^{\frac{1}{3}}V^{\frac{1}{3}}}{2^{-\frac{2}{3}}V^{\frac{1}{3}}} = 2$$

This is the best answer to the question; it tells us that the box with minimum surface area has a base that is twice as wide as its height. This is the optimal shape of the box.

Question: Could we have gotten the answer if we weren't told that the bottom was square?

Answer: Yes. You can do it all in one step if you use multivariable calculus and include variables for both the length and width of the box. Or you can do it in two steps using ideas from this class by first figuring out what rectangle has the smallest perimeter for a fixed area.

Question: Why did you divide x by y?

Answer: We were looking for dimensionless quantities. The lengths x and y are measured in the same units, and x/y gives you the proportions of the box. Another word for what we're interested in here is *proportions*. These are universal, independent of the volume V. The proportions would be the same for any box, at any scale; this is why they provide the nicest answer.

Implicit Differentiation and Min/Max

Example: Find the box (without a top) with least surface area for a fixed volume.

Another way to solve this problem is by using implicit differentiation. As before, this method has some advantages and some disadvantages.

We start the same way:

$$V = x^2 y, \quad A = x^2 + 4xy$$

The goal is to find the minimum value of A while holding V constant.

Next, we just differentiate:

$$\frac{d}{dx} V = 2xy + x^2 \frac{dy}{dx} \Longrightarrow 0 = 2xy + x^2 y'$$

So $y' = -\frac{2y}{x}$.

$$\frac{dA}{dx} = 2x + 4y + 4xy'$$

And when we plug in $y' = -\frac{2y}{x}$ we get:

$$\begin{aligned} \frac{dA}{dx} &= 2x + 4y + 4x\left(-\frac{2y}{x}\right) \\ &= 2x + 4y - 8y \\ \frac{dA}{dx} &= 2x - 4y \end{aligned}$$

To find the critical points, we set $\frac{dA}{dx}$ equal to zero and get $0 = 2x - 4y$ or

$$\frac{x}{y} = 2.$$

This method gets to the answer faster and gets the nicer answer — the scale invariant proportions.

The disadvantage is that we did not check whether this critical point is a maximum, minimum, or neither.

Question: How would we check it?

Answer: By looking at the values of $A(0^+)$ and $A(\infty)$ or perhaps by using your intuition — would a very tall box with a tiny base have more or less surface area than a box that's the lower half of a cube? What about a very short box with a wide base?

Introduction to Related Rates

Next we'll look at the subject of related rates, which will give us another opportunity to practice working with several different variables and equations.

Example: Police are 30 feet from the side of the road. Their radar sees your car approaching at 80 feet per second when your car is 50 feet away from the radar gun. The speed limit is 65 miles per hour (which translates to 95 feet per second). Are you speeding?

How do you set up a problem like this? First, draw a diagram of the setup (as in Fig. 1):

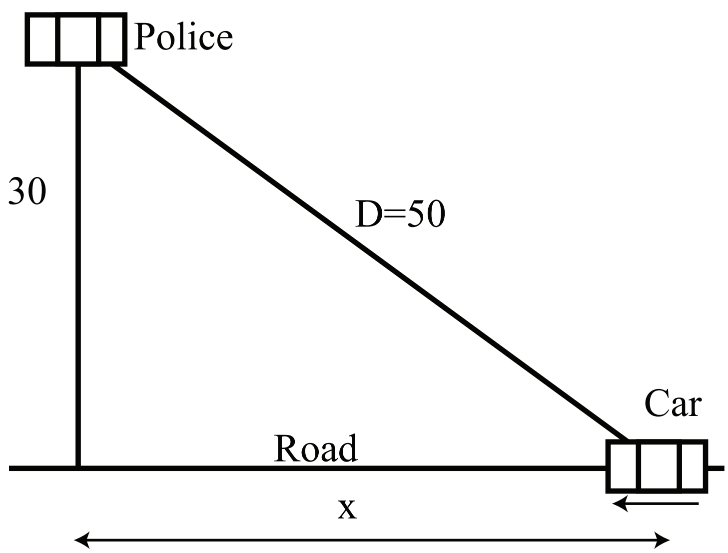


Figure 1: Illustration of example 1: triangle with the police, the car, the road, and labeled variables.

We see that this is a question about a right triangle — the line from the police car to your car is the hypotenuse, the road is one leg of the triangle, and the other leg is a segment from the police car to the road. The Pythagorean Theorem tells us that you're 40 feet away from the right angled vertex of the triangle.

How do we name the variables needed to turn this into a calculus problem? The important thing to figure out is which variables are changing.

We'll choose variable t to stand for time in seconds. The distance along the road from your car to the police car is an important value which we will call x. (We know that $x = 40$ to begin with, but that value will change as the car moves.) Once we have those two variables, we can rewrite the question as "Is $\frac{dx}{dt}$ bigger than 95 feet per second?"

There's another value that is changing, which is the length of the hypotenuse, or the distance between the police car and your car. We'll call this D. (Because we know something about speed limit enforcement, we can correctly assume that the police car is not moving and so the length of the third side of the triangle is not variable.) The rate at which D is changing is exactly what the police are measuring with their radar gun:

$$\frac{dD}{dt} = D' = -80.$$

The derivative must be negative because the value of D is decreasing.

Introduction to Related Rates

We're continuing with a related rates problem from last class.

Example: Police are 30 feet from the side of the road. Their radar sees your car approaching at 80 feet per second when your car is 50 feet away from the radar gun. The speed limit is 65 miles per hour (which translates to 95 feet per second). Are you speeding?

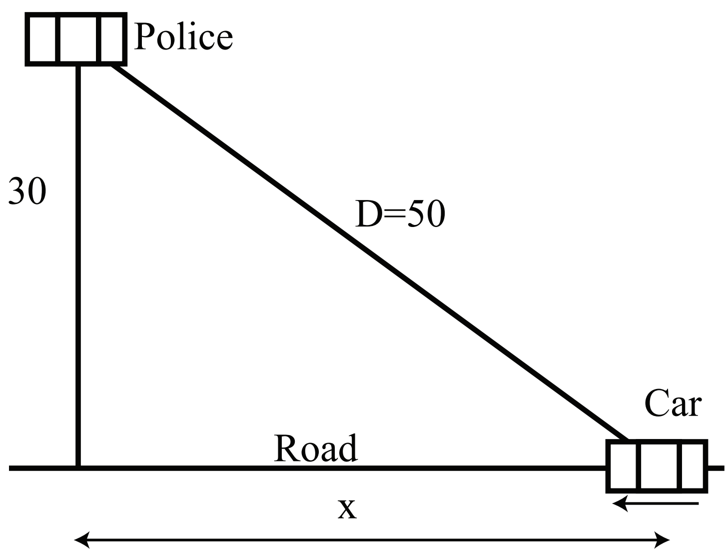


Figure 1: Illustration of example 1: triangle with the police, the car, the road, D and x labeled.

We chose t to stand for time in seconds, x to represent the distance along the road from your car to the police car, and D to represent the straight line distance between your car and the police car.

We know that $\frac{dD}{dt} = -80\text{ft/sec}$ and want to find out whether $\frac{dx}{dt} < -95\text{ft/sec} \cong 65\text{mi/hr}$.

To answer this question we need to understand how x is related to D. First, we know from the Pythagorean theorem that:

$$30^2 + x^2 = D^2$$

We'll differentiate this equation with respect to time using implicit differentiation; we could solve for x, but this would take longer.

While we do this, we must be careful not to replace a variable like D by a constant like 50. The number 50 is a constant — its rate of change is 0. The rate of change of D is -80ft/sec. We must differentiate first before plugging in values.

$$\frac{d}{dt}\left(30^2 + x^2\right) = \frac{d}{dt}\left(D^2\right) \implies 2xx' = 2DD' \implies x' = \frac{2DD'}{2x}$$

Now we can plug in the instantaneous numerical values:

$$x' = \frac{2 \cdot 50 \cdot (-80)}{2 \cdot 40} = -100\frac{\text{feet}}{\text{s}} \cong -68\frac{\text{mi}}{\text{hr}}$$

This exceeds the speed limit of 95 feet per second. You are, in fact, speeding.

There is another, longer, way of solving this problem. Start with:

$$D = \sqrt{30^2 + x^2} = (30^2 + x^2)^{1/2}$$

$$\frac{d}{dt}D = \frac{1}{2}(30^2 + x^2)^{-1/2}(2x\frac{dx}{dt})$$

Plug in the values:

$$-80 = \frac{1}{2}(30^2 + 40^2)^{-1/2}(2)(40)\frac{dx}{dt}$$

and solve to find:

$$\frac{dx}{dt} = -100\frac{\text{feet}}{\text{s}}$$

A third strategy is to differentiate $x = \sqrt{D^2 - 30^2}$. It is easiest to differentiate the equation in its simplest algebraic form $30^2 + x^2 = D^2$, which was our first approach.

The general strategy for these types of problems is:

1. Draw a picture. Set up variables and equations.
2. Take derivatives.
3. Plug in the given values. Don't plug the values in until *after* taking the derivatives.

Related Rates, A Conical Tank

Example: Consider a conical tank whose radius at the top is 4 feet and whose depth is 10 feet. It's being filled with water at the rate of 2 cubic feet per minute. How fast is the water level rising when it is at depth 5 feet?

As always, our first step is to set up a diagram and variables.

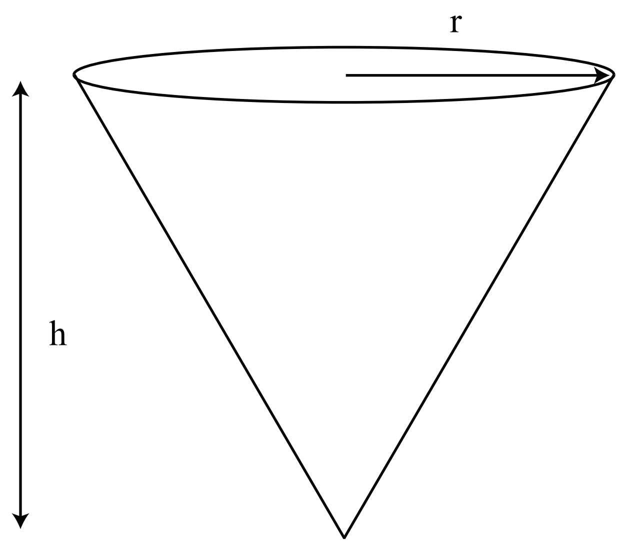


Figure 1: Illustration of example 2: inverted cone water tank.

This diagram just helps us to start thinking about the problem. For instance, we see that because the cone is narrower at the bottom the rate of change of the depth will vary; we need to depict the water level. We also realize that it's difficult to draw useful and accurate diagrams of three dimensional figures — a simple schematic may be more helpful.

The key here is to draw a two-dimensional cross-section. In the figure we're looking at one half of a vertical slice of the tank. The height of the slice equals 10 feet, which is the height of the tank. The widest part of the slice is 4 feet, which is the distance from center to edge of the top of the tank.

We'll use the variable r will represent the distance from center to edge of the top of the water, and h will represent the height of the top of the water (which is also the depth of the water). We can find the relationship between r and h from Fig. 2) using similar triangles:

$$\frac{r}{h} = \frac{4}{10}.$$

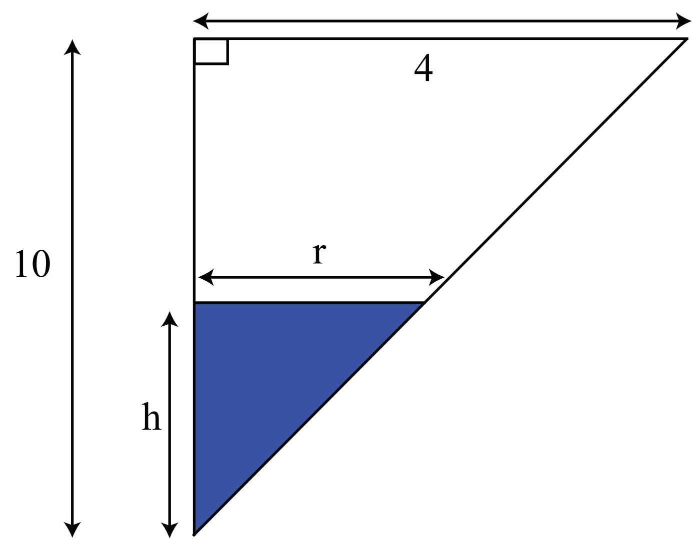


Figure 2: Relating r and h.

Our goal is to find out how fast the water is rising when the tank is half full. What we know is that the volume of water in the tank is changing at a rate of 2 cubic feet per minute. We need equations relating the volume of water in the tank to its depth, h.

The volume of a cone is $\frac{1}{3} \cdot \text{base} \cdot \text{height}$. From Fig. 1), the volume of this tank is given by:

$$V = \frac{1}{3} \cdot \underbrace{\pi r^2}_{\text{base}} \cdot \underbrace{h}_{\text{height}}$$

This relates the volume to the height and radius, and we know the relation between the hight and the radius. We have one more piece of information that we can use: $\frac{dV}{dt} = 2$.

The question is: "What is $\frac{dh}{dt}$ when $h = 5$?"

We've now translated all of the words in the original problem into formulas. Our word problem is now simply a calculus problem.

We could do this by implicit differentiation, but it's easy enough to solve for r in terms of h that there's no need to.

$$r = \frac{2}{5}h.$$

We plug this expression for r back into V to get:

$$V = \frac{1}{3}\pi \left(\frac{2}{5}h\right)^2 h = \frac{4}{3(25)}\pi h^3$$

At this point we could solve for h, but that turns out to be a bad idea. Implicit differentiation is much easier.

$$\begin{aligned}
\frac{dV}{dt} &= \frac{dV}{dh}\frac{dh}{dt} \\
&= \frac{\pi}{3}\left(\frac{2}{5}\right)^2 3h^2\frac{dh}{dt} \\
&= \frac{4}{25}\pi h^2 h'
\end{aligned}$$

Now that we've calculated the rates of change we can plug in the numbers $\frac{dV}{dt} = 2$ and $h = 5$:

$$2 = \left(\frac{4}{25}\right)\pi(5)^2 h'$$

$$2 = 4\pi h'$$

$$h' = \frac{1}{2\pi}\text{ft/min}$$

We were given the rate at which the volume of water in the tank was changing and we used that to compute the rate at which the water in the tank was rising. At the heart of this calculation was the chain rule:

$$\frac{dV}{dt} = \frac{dV}{dh}\frac{dh}{dt}.$$

Related rates problems are all about applying the chain rule to solve word problems.

Ring on a String

We're going to do one more max/min problem.

Consider a ring on a string held fixed at two ends at $(0, 0)$ and (a, b) (see Fig. 1). The ring is free to slide to any point. Find the position (x, y) that the ring slides to.

Note that if $b = 0$, i.e. if the two ends are at equal heights, the ring will settle midway between the two ends ($x = \frac{a}{2}$). We can perform this experiment physically and see the result; we now want to explain that result mathematically. One reason to be interested in this problem is that it's one of many problems that must be solved in order to build a suspension bridge.

Professor Jerison drew a diagram of the possible positions of the ring in lecture by tracing the position of an actual ring on a string held by two students. The next step after drawing this diagram is to name and label the variables, as shown in Figure 1.

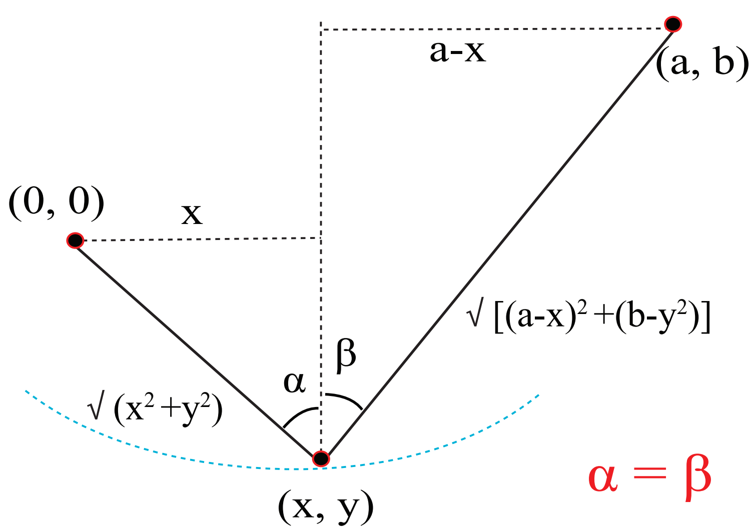


Figure 1: Illustration of the Ring on a String problem.

Physical Principle: The ring settles at the lowest height (lowest potential energy), so the problem is to minimize y subject to the constraint that (x, y) is on the string.

Constraint: The length L of the string is fixed.

$$\sqrt{x^2 + y^2} + \sqrt{(x - a)^2 + (y - b)^2} = L.$$

The function $y = y(x)$ is determined implicitly by the constraint equation above. We traced the constraint curve (possible positions of the ring) on the blackboard; the curve is also suggested in blue in Figure 1. This curve is an ellipse with foci at $(0,0)$ and (a,b), but knowing that the curve is an ellipse does not help us find the lowest point.

Experiments with the hanging ring show that the lowest point is somewhere between $x = 0$ and $x = a$. (This is one way we can confirm that the minimum solution isn't at one of the ends of the string; don't try to use the second derivative test.) Since the ends of the constraint curve are higher than the middle, the lowest point is a critical point (a point where $y'(x) = 0$). In class we also gave a physical demonstration of this by drawing the horizontal tangent at the lowest point.

To find the critical point, differentiate the constraint equation implicitly with respect to x:

$$\frac{x + yy'}{\sqrt{x^2+y^2}} + \frac{x - a + (y-b)y'}{\sqrt{(x-a)^2+(y-b)^2}} = 0.$$

Since $y' = 0$ a the critical point, the equation can be rewritten as:

$$\frac{x}{\sqrt{x^2+y^2}} = \frac{a-x}{\sqrt{(x-a)^2+(y-b)^2}}$$

From Fig. 1, we see that the last equation can be interpreted geometrically as saying that:

$$\sin\alpha = \sin\beta \quad \Longrightarrow \quad \alpha = \beta,$$

where α and β are the angles the left and right portions of the string make with the vertical.

Physical and geometric conclusions

The angles α and β are equal.

Using vectors to compute the force exerted by gravity on the two halves of the string, one finds that there is *equal tension* in the two halves of the string — a physical equilibrium. This is desirable in construction; if one end is under more stress than the other, it's more likely to break.

From another point of view, the equal angle property expresses a geometric property of ellipses: Suppose that the ellipse is a mirror. A ray of light from the focus $(0,0)$ reflects off the mirror according to the rule angle of incidence equals angle of reflection, and therefore the ray goes directly to the other focus at (a,b). This was used to good effect in the "Strokes of Genius: Mini Golf by Artists" exhibit at the DeCordova museum in the early 1990's; by placing the tee at one focus of an ellipse and the hole at the other, an artist created a golf course on which any stroke would end with a hole in one.

Formulae for x and y

We did not yet find the location of (x,y). We will now show that:

$$x = \frac{a}{2}\left(1 - \frac{b}{\sqrt{L^2-a^2}}\right), \quad y = \frac{1}{2}\left(b - \sqrt{L^2-a^2}\right).$$

Because $\alpha = \beta$,

$$x = \sqrt{x^2 + y^2}\ \sin\alpha; \quad a - x = \sqrt{(x-a)^2 + (y-b)^2}\ \sin\alpha$$

Adding these two equations,

$$a = \left(\sqrt{x^2+y^2} + \sqrt{(x-a)^2 + (y-b)^2}\right)\sin\alpha = L\sin\alpha \implies \sin\alpha = \frac{a}{L}$$

The equations for the vertical legs of the right triangles are (note that $y < 0$):

$$-y = \sqrt{x^2 + y^2}\ \cos\alpha; \quad b - y = \sqrt{(x-a)^2 + (y-b)^2}\ \cos\beta.$$

Adding these two equations, and using $\alpha = \beta$, we get:

$$b - 2y = \left(\sqrt{x^2+y^2} + \sqrt{(x-a)^2 + (y-b)^2}\right)\cos\alpha = L\cos\alpha$$

$$\implies$$

$$y = \frac{1}{2}(b - L\cos\alpha).$$

Use the relation $\sin\alpha = \dfrac{a}{L}$ to write:

$$\begin{aligned} L\cos\alpha &= L\sqrt{1-\sin^2\alpha} \\ &= \sqrt{L^2 - a^2}. \end{aligned}$$

Then the formula for y is:

$$y = \frac{1}{2}\left(b - \sqrt{L^2 - a^2}\right).$$

Finally, to find the formula for x, use similar right triangles:

$$\tan\alpha = \frac{x}{-y} = \frac{a-x}{b-y} \implies x(b-y) = (-y)(a-x) \implies (b-2y)x = -ay$$

Therefore,

$$x = \frac{-ay}{b-2y} = \frac{a}{2}\left(1 - \frac{b}{\sqrt{L^2-a^2}}\right).$$

Thus we have formulae for x and y in terms of a, b and L.

This derivation of the formulae for x and y wasn't covered in lecture because it is long and because the most illuminating part of the problem is the balance condition $\alpha = \beta$ that is an immediate consequence of the critical point computation.

Final Remark. In 18.02, you will learn to treat constrained max/min problems in any number of variables using a method called Lagrange multipliers.

Newton's Method

Newton's method is a powerful tool for solving equations of the form $f(x) = 0$.

Example: Solve $x^2 = 5$.

We're going to use Newton's method to find a numerical approximation for $\sqrt{5}$. Any equation that you understand can be solved this way. In order to use Newton's method, we define $f(x) = x^2 - 5$. By finding the value of x for which $f(x) = 0$ we solve the equation $x^2 = 5$.

Our goal is to discover where the graph crosses the x-axis. We start with an initial guess — we'll guess $x_0 = 2$, since $\sqrt{5} \approx \sqrt{4} = 2$. This is not a very good guess; $f(2) = -1$, and we're looking for a number x for which $f(x) = 0$. We'll try to improve our guess.

We pretend that the function is linear, and look for the point where the tangent line to the function at x_0 crosses the x-axis: see Fig. 1. This point $(x_1, 0)$ gives us a new guess at our solution: x_1.

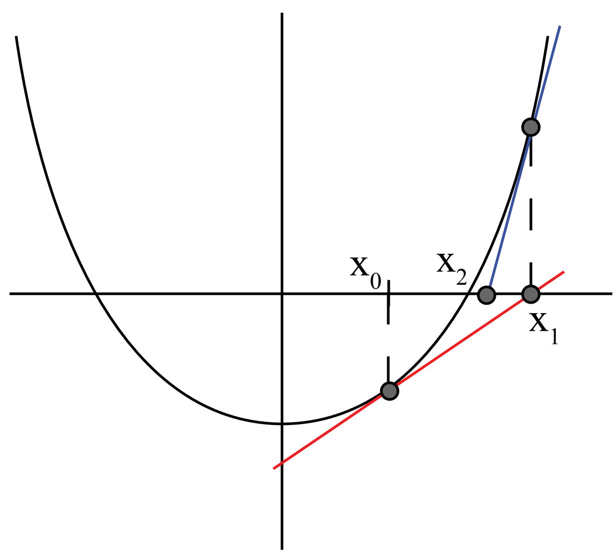


Figure 1: Illustration of Newton's Method

The equation for the tangent line is:

$$y - y_0 = m(x - x_0)$$

When the tangent line intercepts the x-axis $y = 0$, and the x coordinate of that point is our new guess x_1.

$$\begin{aligned} -y_0 &= m(x_1 - x_0) \\ -\frac{y_0}{m} &= x_1 - x_0 \\ x_1 &= x_0 - \frac{y_0}{m} \end{aligned}$$

In terms of f:

$$\begin{aligned} y_0 &= f(x_0) \\ m &= f'(x_0) \end{aligned}$$

because m is the slope of the tangent line to $y = f(x)$ at the point (x_0, y_0). Therefore,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

The point of Newton's method is that we can improve our new guess by repeating this process. To get our $(n+1)^{st}$ guess we apply this formula to our n^{th} guess:

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.}$$

In our example, $x_0 = 2$ and $f(x) = x^2 - 5$. We first calculate $f'(x) = 2x$. Thus,

$$\begin{aligned} x_1 &= x_0 - \frac{(x_0^2 - 5)}{2x_0} = x_0 - \frac{1}{2}x_0 + \frac{5}{2x_0} \\ x_1 &= \frac{1}{2}x_0 + \frac{5}{2x_0} \end{aligned}$$

The main idea is to repeat (iterate) this process:

$$\begin{aligned} x_2 &= \frac{1}{2}x_1 + \frac{5}{2x_1} \\ x_3 &= \frac{1}{2}x_2 + \frac{5}{2x_2} \end{aligned}$$

and so on. The procedure approximates $\sqrt{5}$ extremely well.

Let's see how well this works:

$$\begin{aligned}
x_1 &= \frac{1}{2}2 + \frac{5}{2 \cdot 2} \\
&= 1 + \frac{5}{4} \\
&= \frac{9}{4} \\
x_2 &= \frac{1}{2}\frac{9}{4} + \frac{5}{2\frac{9}{4}} \\
&= \frac{9}{8} + \frac{5}{2}\frac{4}{9} \\
&= \frac{9}{8} + \frac{10}{9} \\
&= \frac{161}{72} \\
x_3 &= \frac{1}{2}\frac{161}{72} + \frac{5}{2}\frac{72}{161}
\end{aligned}$$

n	x_n	$\sqrt{5} - x_n$
0	2	2×10^{-1}
1	$\frac{9}{4}$	10^{-2}
2	$\frac{161}{72}$	4×10^{-5}
3	$\frac{1}{2}\frac{161}{72} + \frac{5}{2}\frac{72}{161}$	10^{-10}

Notice that the number of digits of accuracy doubles with each iteration; x_2 is as good an approximation as you'll ever need, and x_3 is as good an approximation as the one displayed by your calculator.

Newton's Method

Today we'll discuss the accuracy of Newton's Method.

Recall how Newton's method works: to find the point at which a graph crosses the x-axis you make an initial guess x_0 at the x-coordinate of that crossing. You then find the tangent line to the graph at x_0 and use it to improve your guess: x_1 is the x-coordinate at which the tangent line crosses the x-axis. (See Fig. 1.) You can now draw the tangent line at x_1 to get a new guess x_2, and so on.

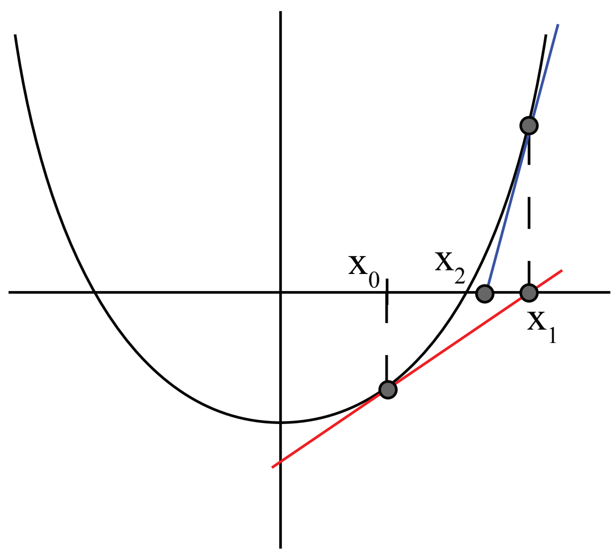


Figure 1: Illustration of Newton's Method

In algebraic terms,

$$\boxed{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}}$$

Figure 2 illustrates the k^{th} iteration of Newton's method.

If we're going to use this to get numerical approximations of solutions, we should know how accurate it is. If x is the exact value of the solution, then x_1 is $E_1 = |x - x_1|$ away from the exact answer. The error in our approximation at step n is $E_n = |x - x_n|$.

Last time we saw that error values of $E_n = |\sqrt{5} - x_n|$ quickly became very close to zero. It turns out that $E_2 \sim E_1^2$. So if $E_0 = 10^-1$, the size of the error can be expected to decrease as follows:

E_0	E_1	E_2	E_3	E_4
10^{-1}	10^{-2}	10^{-4}	10^{-8}	10^{-16}

The number of digits of accuracy doubles at each step!

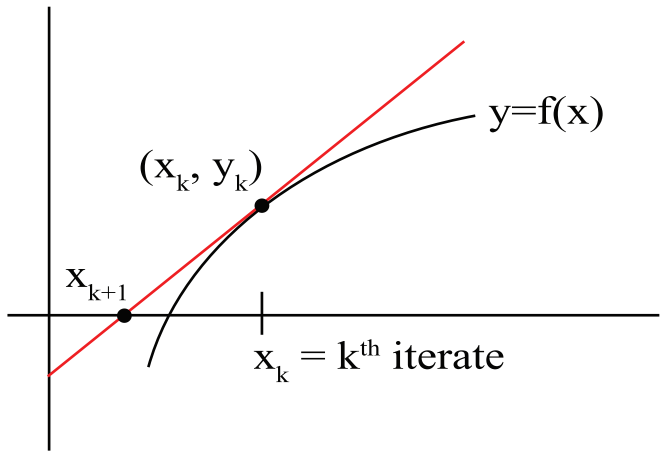


Figure 2: Illustration of Newton's Method.

Newton's method works (very) well if $|f'|$ is not too small, $|f''|$ is not too big, and x_0 starts near the solution x.

Newton's Method: What Could Go Wrong?

Newton's method works (very) well if $|f'|$ is not too small, $|f''|$ is not too big, and x_0 starts near the solution x.

We're not going to discuss these conditions in detail, but let's see why they're there. If f'' is too large the graph would be sharply curved, in which case the tangent line might not be a good approximation to the graph and x_1 might not be close to the solution. There are a couple of things that can go wrong if x_0 is too far from x_1, which we'll discuss now.

If the error $E_0 = |x - x_0|$ is greater than 1 and $E_1 \sim E_0^2$, the error of your estimate could actually *increase* as you apply Newton's method.

In the example $f(x) = x^2 - 5$, if we had chosen $x_0 = -2$ we would have found the solution $-\sqrt{5}$ and not $\sqrt{5}$. This convergence to an unexpected root is illustrated in Fig. 1

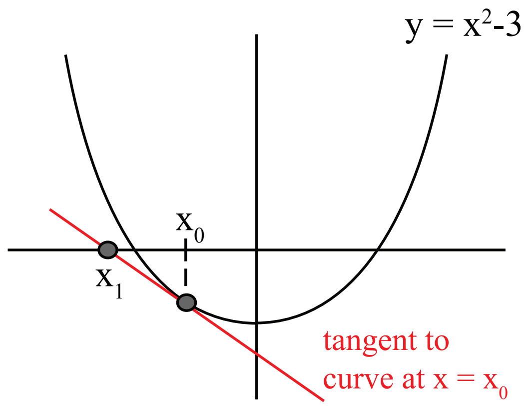


Figure 1: Newton's method converging to an unexpected root.

In the same example, if we chose $x_0 = 0$ then $f'(x_0) = 0$ and $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$ is undefined.

Finally, there's a chance that Newton's method will cycle back and forth between two value and never converge at all. This failure is illustrated in Fig. 2; $x_2 = x_0$, $x_3 = x_1$, and so forth.

Newton's method is a good way of approximating solutions, but applying it requires some intelligence. You must beware of getting an unexpected result or no result at all. The better your initial guess at the solution, the more likely you are to get a correct result.

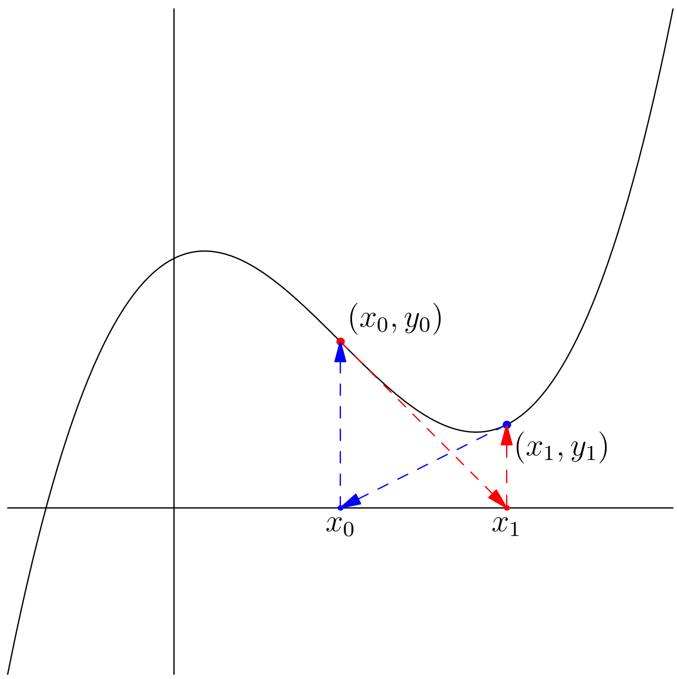


Figure 2: Newton's method cycling between x_0 and x_1.

The Mean Value Theorem

The mean value theorem is a little theoretical, and will allow us to introduce the idea of integration in a few lectures. Integration is the subject of the second half of this course. We'll use the abbreviation "MVT" when discussing it.

Colloquially, the MVT theorem tells you that if you fly 3,000 kilometers in 6 hours, at some time during the flight you will be traveling at a speed of 500 kilometers per hour. (Because your average speed is 500 km/hr.)

The reason it's called the "mean value theorem" is because the word "mean" is the same as the word "average".

In math symbols, it says:

$$\boxed{\begin{array}{c} \dfrac{f(b)-f(a)}{b-a} = f'(c) \quad \text{(for some } c,\ a < c < b) \\ \text{Provided that } f \text{ is differentiable on } a < x < b, \text{ and continuous on } a \le x \le b. \end{array}}$$

Geometric Proof of MVT: Consider the graph of $f(x)$. Here, $\dfrac{f(b)-f(a)}{b-a}$ is the slope of a secant line joining the points $(a, f(a))$ and $(b, f(b))$, and $f'(c)$ is the slope of a tangent line. We need to show that somewhere between a and b there's a point on the graph $(c, f(c))$ whose tangent line has the same slope as that secant line.

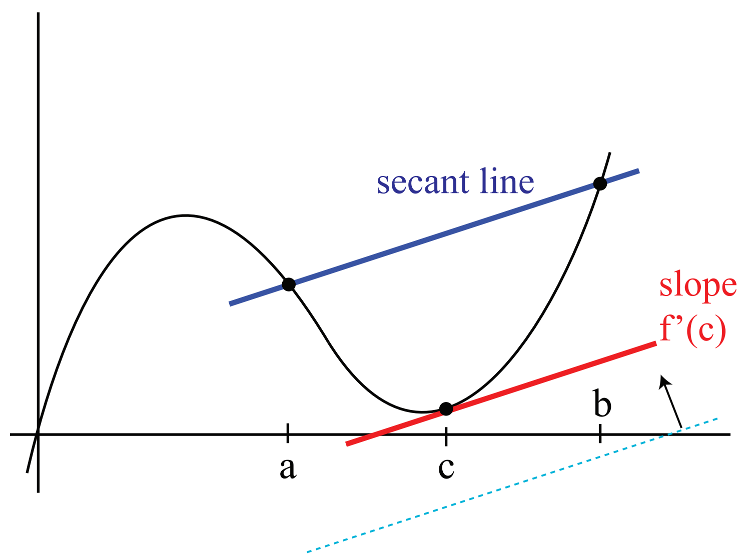


Figure 1: Illustration of the Mean Value Theorem.

Take (dotted) lines parallel to the secant line, as in Fig. 1 and shift them up from below the graph until one of them first touches the portion of the graph that lies between a and b. If it does not touch, start with a dotted line above the graph and move it down until it touches.

When reading a proof, you should always be thinking about why the hypotheses are necessary. Would the proof still work if the function were discontinuous or if it were not differentiable?

We need the hypotheses that f is continuous, because if you could sit still for six hours and then instantly teleport 3,000 km there would never be a time at which you were traveling 500 km/hour. The mean value theorem can't make any guarantees about discontinuous functions.

What if the function isn't differentiable? Suppose $f(x) = |x|$. Then the dotted line always touches the graph first at $x = 0$, no matter what its slope is (see Fig. 2). Even though f is differentiable everywhere except $x = 0$, the mean value theorem still doesn't work here; we need $f'(x)$ to exist at *all* x between a and b.

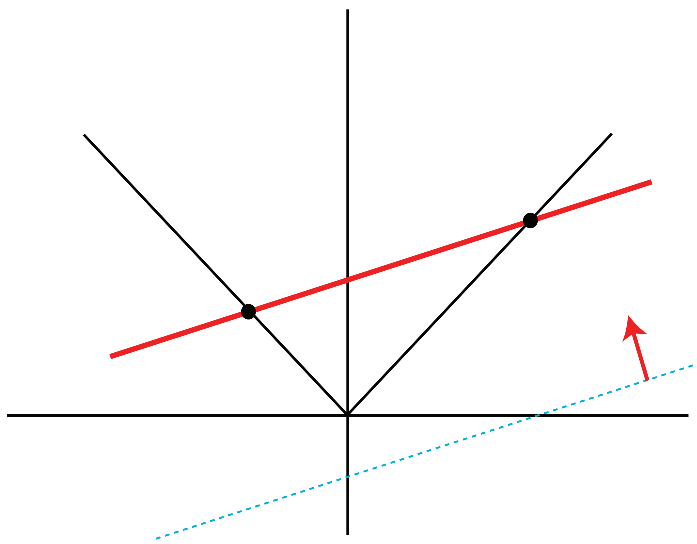


Figure 2: Graph of $y = |x|$ with secant line. (One bad point ruins the proof.)

Question: What if the line parallel to the secant line touches the graph in more than one point?

Answer: The more the merrier! The graph could wiggle a lot of times and the line could touch in ten places, or f could be constant and the line could touch every point on the graph at once. In mathematics, when we claim something is true for one point we don't necessarily mean that it isn't true for others.

The fact that this point exists is a touchy point; we can see why it ought to exist but we didn't really prove that it does. The formal proof has to do with the existence of tangent lines and uses more analysis then we can do in this class.

Mean Value Theorem: Consequences

The first thing we apply the MVT to is graphing, but we'll see later that this is significant in all the rest of calculus.

- If $f' > 0$ then f is increasing.
- If $f' < 0$ then f is decreasing.
- If $f' = 0$ then f is constant.

We told you that the first of these two are true, but we didn't prove them. We can now prove them using the MVT.

Proof: The mean value theorem tells us that:

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

for some c between a and b. For the purposes of this proof we'll assume that $b > a$. We write the equation for the MVT "backwards" because we want to use information about f' to get information about f.

We manipulate the equation to get:

$$\begin{aligned} f(b) - f(a) &= f'(c)(b - a) \\ f(b) &= f(a) + f'(c)(b - a) \end{aligned}$$

This new form of the MVT will let us check these three facts.

Since $a < b$, $b - a > 0$ and the sign of $f'(c)(b - a)$ is completely determined by the sign of $f'(c)(b - a)$.

- If $f'(c) > 0$ then $f(b) > f(a)$.
- If $f'(c) < 0$ then $f(b) < f(a)$.
- If $f'(c) = 0$ then $f(b) = f(a)$.

These facts may seem obvious, but they are not. The definition of the derivative is written in terms of infinitesimals. It's not a sure thing that these infinitesimals have anything to do with the large scale behavior of the function. Before, we were saying that the difference quotient was approximately equal to the derivative. Now we're saying that it's exactly equal to a derivative. (Although we don't know at what point that derivative should be taken.)

The Mean Value Theorem and Linear Approximation

What's the difference between the mean value theorem and the linear approximation?

The linear approximation to $f(x)$ near a has the formula:

$$f(x) \approx f(a) + f'(a)(x - a) \quad x \text{ near } a.$$

If we let $\Delta x = x - a$, we get:

$$\begin{aligned} f(x) &\approx f(a) + f'(a)(x - a) \\ f(x) - f(a) &\approx f'(a)\Delta x \\ \frac{\Delta f}{\Delta x} &\approx f'(a). \end{aligned}$$

Similarly the MVT says:

$$f(b) = f(a) + f'(c)(b - a) \quad \text{for some } c, a < c < b$$

If b is near a then we can write $b - a = \Delta x$ and rewrite the theorem as:

$$\frac{\Delta f}{\Delta x} = f'(c) \quad \text{for some } c, a < c < b.$$

The mean value theorem tells us that $\frac{\Delta f}{\Delta x}$ is *exactly* equal to $f'(c)$ for some c between a and b. We don't know precisely where c is; it depends on f, a, and b.

As Professor Jerison says in the video, this is telling us that the average change on the interval is between the maximum and minimum values $f'(x)$ reaches on the interval $[a, b]$ (because the derivative is continuous).

$$\min_{a \leq x \leq b} f'(x) \leq \frac{f(b) - f(a)}{b - a} = f'(c) \leq \max_{a \leq x \leq b} f'(x)$$

In other words, the average speed of your trip is somewhere between your minimum speed and your maximum speed.

Linear approximation, is based on the assumption that the average speed is approximately equal to the initial (or possibly final) speed. Figure 1 illustrates the approximation $1 + x \approx e^x$.

If the interval $[a, b]$ is short, $f'(x)$ won't vary much between a and b; the max and the min should be pretty close. The mean value theorem tells us absolutely that the slope of the secant line from $(a, f(a))$ to $(x, f(x))$ is no less than the minimum value and no more than the maximum value of f' on that interval, which assures us that the linear approximation does give us a reasonable approximation of the f.

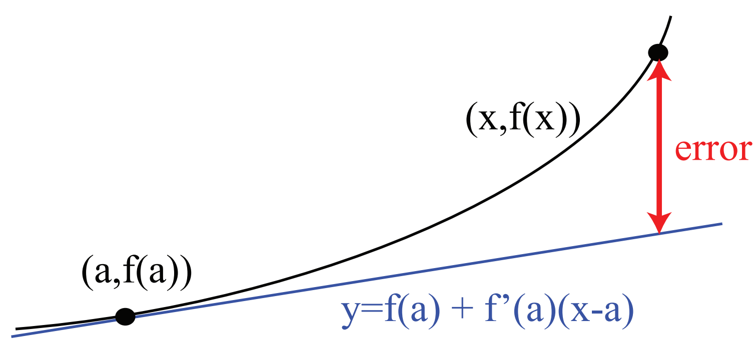


Figure 1: MVT vs. Linear Approximation.

The Mean Value Theorem and Inequalities

The mean value theorem tells us that if f and f' are continuous on $[a, b]$ then:

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

for some value c between a and b. Since f' is continuous, $f'(c)$ must lie between the minimum and maximum values of $f'(x)$ on $[a, b]$. In other words:

$$\min_{a \leq x \leq b} f'(x) \leq \frac{f(b) - f(a)}{b - a} = f'(c) \leq \max_{a \leq x \leq b} f'(x).$$

This is the form that the mean value theorem takes when it is used in problem solving (as opposed to mathematical proofs), and this is the form that you will need to know for the test.

In practice, you may even forget the mean value theorem and remember only these three inequalities:

- If $f'(c) > 0$ then $f(b) > f(a)$.
- If $f'(c) < 0$ then $f(b) < f(a)$.
- If $f'(c) = 0$ then $f(b) = f(a)$.

These can be used to prove mathematical inequalities. The following examples compare the function e^x to its linear and quadratic approximations and are the first steps toward a deeper understanding of the function.

Example: Show that $e^x > 1 + x$ for $x > 0$.

To prove this, we'll instead show that $f(x) = e^x - (1 + x)$ is always positive. We know that $f(0) = e^0 - (1 + 0) = 0$ and $f'(x) = e^x - 1$. When x is positive, $f'(x)$ is positive because $e^x > 1$.

We know that if $f'(x) > 0$ on an interval then $f(x)$ is increasing on that interval, so we can conclude that $f(x) > f(0)$ for $x > 0$. In other words,

$$e^x - (1 + x) > 0 \iff e^x > 1 + x.$$

Example: Show that $e^x > 1 + x + \frac{x^2}{2}$ for $x > 0$.

The value of $1 + x + \frac{x^2}{2}$ is slightly greater than that of $1 + x$, but it turns out that it's still less than the value of e^x. We let $g(x) = e^x - (1 + x + \frac{x^2}{2})$ and do the same thing we did before:

$$\begin{aligned} g(0) &= 1 - (1) = 0 \\ g'(x) &= e^x - (1 + x) \end{aligned}$$

We know $g'(x) > 0$ because we proved $f(x) > 0$ in the above example. Since $g'(x)$ is positive, g is increasing for $x > 0$, so $g(x) > g(0)$ when $x > 0$, so $e^x - (1 + x + \frac{x^2}{2}) > 0$ and $e^x > (1 + x + \frac{x^2}{2})$.

We can keep on going: $e^x > 1 + x + \dfrac{x^2}{2} + \dfrac{x^3}{3!}$ for $x > 0$. Eventually, it turns out that

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots \quad \text{(an infinite sum)}$$

We will be discussing this when we get to Taylor series near the end of the course.

Differentials

Today we move on from differentiation to integration. For this we'll need a new notation for quantities called differentials.

Given a function $y = f(x)$, the *differential* of y is

$$\boxed{dy = f'(x)dx}$$

Because $y = f(x)$ we sometimes call this the differential of f. Both dy and $f'(x)dx$ are called *differentials.* You can think of

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

as a quotient of differentials. Get used to this idea; it comes up in many contexts, including this class and multivariable calculus.

This arises from the Leibniz interpretation of a derivative as a ratio of "infinitesimal" quantities; differentials are sort of like infinitely small quantities.

Working with differentials is much more effective than using the notation coined by Newton; good notation can help you think much faster. Leibniz's notation was adopted on the Continent and Newton dominated in Britain; as a result the British fell behind by one or two hundred years in the development of calculus.

Differentials and Linear Approximation

Linear approximation allows us to estimate the value of $f(x+\Delta x)$ based on the values of $f(x)$ and $f'(x)$. We replace the change in horizontal position Δx by the differential dx. Similarly, we replace the change in height Δy by dy. (See Figure 1.)

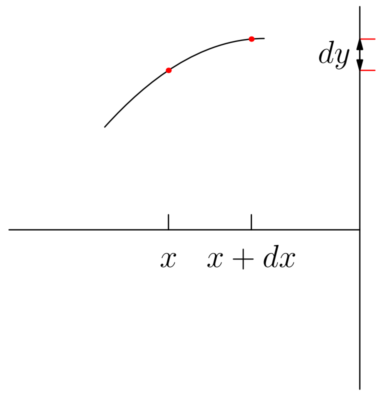


Figure 1: We use dx and dy in place of Δx and Δy.

Example: Find the approximate value of $(64.1)^{\frac{1}{3}}$.

Method 1 (using differentials)

We're going to use a linear approximation of the function $y = f(x) = x^{\frac{1}{3}}$. Our base point will be $x_0 = 64$ because it's easy to compute $y_0 = 64^{\frac{1}{3}} = 4$. By definition, $dy = f'(x)dx = \frac{1}{3}x^{-\frac{2}{3}}dx$.

$$\begin{aligned} dy &= \frac{1}{3}(64)^{-\frac{2}{3}}dx \\ &= \frac{1}{3}\frac{1}{16}dx \\ &= \frac{1}{48}dx \end{aligned}$$

We want to approximate $(64.1)^{\frac{1}{3}}$, so $x + dx = 64.1$ and $dx = 0.1 = \frac{1}{10}$. At the value $64.1 = x_0 + dx$, $f(x)$ is exactly equal to $y_0 + \Delta y$ (because this is how we defined Δy) and is approximately equal to $y_0 + dy$, where dy is is linear in dx as derived above.

In essence, the point $(x_0 + dx, y_0 + dy)$ is an infinitesimally small step away from (x_0, y_0) along the tangent line. Of course $\frac{1}{10}$ is not infinitesimally small, which is why this is an approximation rather than an exact value.

$$(64.1)^{\frac{1}{3}} \approx y + dy$$

$$\begin{aligned} &\approx 4 + \frac{1}{48}dx \\ &\approx 4 + \frac{1}{48}\frac{1}{10} \\ &\approx 4.002 \end{aligned}$$

Method 2 (review)

When we compare this to our previous notation we discover that the calculations are the same; only the notation has changed.

The basic formula for linear approximation is:

$$f(x) = f(a) + f'(a)(x - a)$$

Here $a = 64$ and $f(x) = x^{\frac{1}{3}}$, so $f(a) = f(64) = 4$ and $f'(a) = \frac{1}{3}a^{-\frac{2}{3}} = \frac{1}{48}$

Our approximation then becomes:

$$\begin{aligned} f(x) &\approx f(a) + f'(a)(x - a) \\ x^{\frac{1}{3}} &\approx 4 + \frac{1}{48}(x - 64) \\ (64.1)^{\frac{1}{3}} &\approx 4 + \frac{1}{48}\frac{1}{10} \\ (64.1)^{\frac{1}{3}} &\approx 4.002 \end{aligned}$$

We get the same answer as before, by doing a nearly identical calculation.

Introduction to Antiderivatives

This is a new notation and also a new concept. $G(x) = \displaystyle\int g(x)dx$ is the *antiderivative* of g. Other ways of saying this are:

$$G'(x) = g(x) \quad \text{or,} \quad dG = g(x)dx$$

There are a few things to notice about this definition. It includes a differential dx. It also includes the symbol $\int$, called an *integral sign*; the expression $\displaystyle\int g(x)dx$ is an *integral*. Another name for the antiderivative of g is the *indefinite integral* of g. (We'll learn what "indefinite" means in this context very shortly.)

If $G(x)$ is the antiderivative of $g(x)$ then $G'(x) = g(x)$. To find the antiderivative of a function g (to integrate g), we need to find a function whose derivative is g. In practice, finding antiderivatives is not as easy as finding derivatives, but we want to be able to integrate as many things as possible. We'll start with some examples.

Example: $\sin x$

We start with the integral of $g(x) = \sin x$. This is a function whose derivative is $\sin x$. What function has $\sin x$ as its derivative?

Student: $-\cos x$

Because the derivative of $-\cos x$ is $\sin x$, this is an antiderivative of $\sin x$. If:

$$\begin{aligned} G(x) &= -\cos x, \quad \text{then} \\ G'(x) &= \sin x \end{aligned}$$

On the other hand, if we had instead chosen $G(x) = -\cos x + 7$ we would still have had $G'(x) = \sin x$. Because the derivative of a constant is 0, we can add any constant to $G(x)$ and still have an antiderivative of $\sin x$. We write:

$$\int \sin x \; dx = -\cos x + c$$

and call this the *indefinite integral* of $\sin x$ because c can be any constant — it's an indefinite value. Whenever we take the antiderivative of something our answer is ambiguous up to a constant.

Antiderivative of x^a

What function has the derivative x^a? We know that the exponent decreases by one when we differentiate, so we guess x^{1+1}. This doesn't quite work:

$$d(x^{a+1}) = (a+1)x^a dx.$$

We have to divide both sides by the constant $(a+1)$ to get the correct answer.

$$\begin{aligned} d(\frac{x^{a+1}}{a+1}) &= x^a dx \\ \frac{x^{a+1}}{a+1} + c &= \int x^a dx \end{aligned}$$

But wait! Although it's true that $d(x^{a+1}) = (a+1)x^a dx$, it is not always true that $\int x^a dx = \frac{x^{a+1}}{a+1} + c$. When $a = -1$ the denominator is zero. However, we can still say that $\int x^a dx = \frac{x^{a+1}}{a+1} + c$ for $a \neq -1$.

What happens when $a = -1$? What is $\int \frac{1}{x} dx$?

So far we've used the formulas $\frac{d}{dx} \cos x = -\sin x$ and $\frac{d}{dx} x^{n+1} = (n+1)x^n$. An important part of integration is remembering formulas for derivatives and "reading them backward". In this case, the formula we need is $\frac{d}{dx} \ln x = \frac{1}{x}$. Using this, we get $\int \frac{1}{x} dx = \ln x + c$.

This formula is fine when $x > 0$, but $\ln x$ is not defined when x is negative. The more standard form of this equation is:

$$\int \frac{1}{x} dx = \ln |x| + c.$$

The absolute value doesn't change anything when $x \geq 0$, so we only need to check this formula when x is negative. In order to do so, we have to differentiate $\ln |x|$.

$$\begin{aligned} \frac{d}{dx} \ln |x| &= \frac{d}{dx} \ln(-x) \quad (|x| = -x \text{ when } x < 0) \\ &= \frac{1}{-x} \frac{d}{dx}(-x) \quad \text{(by the chain rule)} \\ &= -\frac{1}{-x} \\ &= \frac{1}{x} \end{aligned}$$

If we graph $\ln |x|$ we can see that this function does have slope $\frac{1}{x}$.

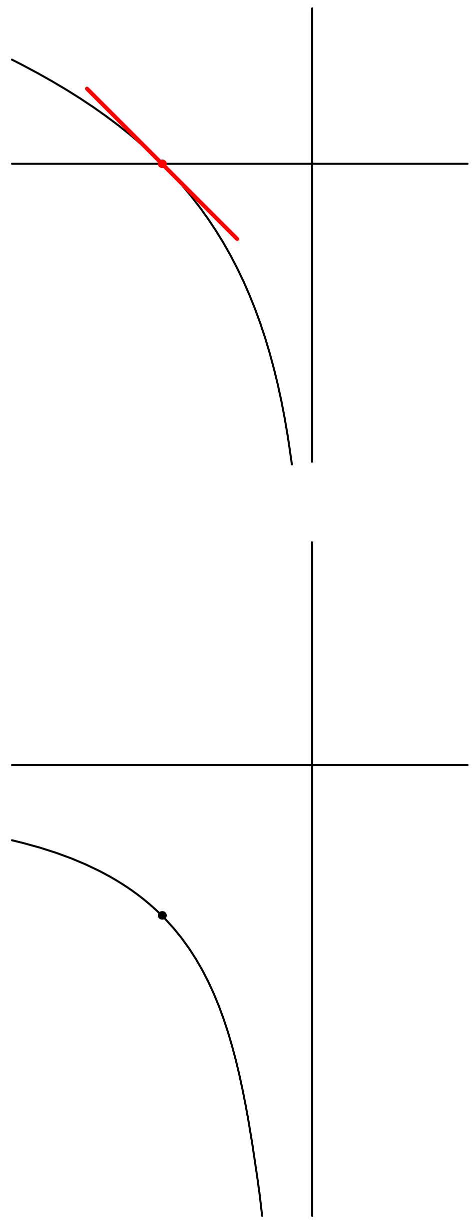


Figure 1: Graphs of $y = ln(-x)$ (above) and $y' = \dfrac{1}{x}$ (below).

Antiderivatives of $\sec^2 x$ and $\dfrac{1}{\sqrt{1-x^2}}$

Example: $\displaystyle\int \sec^2 x\,dx$

Searching for antiderivatives will help you remember the specific formulas for derivatives. In this case, you need to remember that $\frac{d}{dx}\tan x = \sec^2 x$.

$$\int \sec^2 x\,dx = \tan x + c$$

Example: $\displaystyle\int \frac{1}{\sqrt{1-x^2}}\,dx$

An alternate way to write this integral is $\int \frac{dx}{\sqrt{1-x^2}}$. This is consistent with the idea that dx is an infinitesimal quantity which can be treated like any other number.

We remember $\frac{d}{dx}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$ and conclude that:

$$\int \frac{dx}{\sqrt{1-x^2}} = \sin^{-1} x + c.$$

Example: $\displaystyle\int \frac{dx}{1+x^2}$

$$\int \frac{dx}{1+x^2} = \tan^{-1} x + c$$

When looking for antiderivatives, you'll spend a lot of time thinking about derivatives. For a little while you may get the two mixed up and differentiate where you were meant to integrate, or vice-versa. With practice, this problem goes away.

Here is a list of the antiderivatives presented in this lecture:

1. $\displaystyle\int \sin x\,dx = -\cos x + c$ where c is any constant.
2. $\displaystyle\int x^n\,dx = \frac{x^{n+1}}{n+1} + c$ for $n \neq -1$.
3. $\displaystyle\int \frac{dx}{x} = \ln|x| + c$ (This takes care of the exceptional case $n = -1$ in 2.)
4. $\displaystyle\int \sec^2 x\,dx = \tan x + c$.
5. $\displaystyle\int \frac{dx}{\sqrt{1-x^2}} = \sin^{-1} x + c$ (where $\sin^{-1} x$ denotes "inverse sine" or arcsin, and not $\dfrac{1}{\sin x}$.)
6. $\displaystyle\int \frac{dx}{1+x^2} = \tan^{-1}(x) + c$.

Antiderivatives are Unique up to a Constant

Theorem: If $F'(x) = f(x)$ and $G'(x) = f(x)$, then $F(x) = G(x) + c$.

In other words, once we've found one antiderivative of a function we know that any other antidervative we might find will only differ from it by some added constant.

Proof: If $F' = G'$ then $(F - G)' = F' - G' = f - f = 0$.

Recall that we proved as a corollary of the Mean Value Theorem that if a function's derivative is zero then it is constant. Hence $G(x) - F(x) = c$ (for some constant c). That is, $G(x) = F(x) + c$.

This is a very important fact. It's the basis for calculus; the reason why it makes sense to do calculus at all. This theorem tells us that if we know the rate of change of a function we can find out everything else about the function except this starting value c.

Substitution: $\int x^3(x^4+2)^5 dx$

We want to compute $\int x^3(x^4+2)^5 dx$.

We already have a formula for $\int x^n dx$, so we could expand $(x^4+2)^5$ and integrate the polynomial. That would be messy. Instead we'll use the method of substitution.

Finding the exact integral of a function is much harder than finding its derivative; occasionally it's impossible. In this unit we're only going to use one method, which means that whenever you see an integral, either you'll be able to divine immediately what the answer is or you'll use substitution. The method of substitution is tailor-made for differential notation.

To use the method of substitution, find the messiest function in your integral; in this case that will be $u = x^4 + 2$. The differential of u is $du = u'dx = 4x^3dx$. Luckily, we can substitute these two expressions into our original integral and simplify it considerably.

The original problem was to find $\int x^3(x^4+2)^5 dx$. We can replace $(x^4+2)^5$ by u^5, and $x^3dx = \frac{1}{4}du$:

$$\begin{aligned} \int x^3(x^4+2)^5 dx &= \int \underbrace{(x^4+2)^5}_{u^5}\underbrace{x^3dx}_{\frac{1}{4}du} \\ &= \int \frac{u^5du}{4} \\ &= \frac{1}{4}\cdot\frac{1}{6}u^6 = \frac{u^6}{24} \end{aligned}$$

This is not the answer to the question, because this answer is expressed in terms of u. The problem was posed in terms of the variable x. We change variables back to x by plugging in our definition of u:

$$\frac{u^6}{24} = \frac{(x^4+2)^6}{24}.$$

We conclude that:

$$\int x^3(x^4+2)^5 dx = \frac{(x^4+2)^6}{24}.$$

Integration by "Advanced Guessing"

Example: $\displaystyle\int \frac{xdx}{\sqrt{1+x^2}}$

If we use the method of substitution, we start by setting u equal to the ugliest part of our integral:

$$u = 1 + x^2 \quad \text{and} \quad du = 2xdx.$$

The calculation looks like:

$$\begin{aligned}
\int \frac{xdx}{\sqrt{1+x^2}} &= \int \frac{\frac{1}{2}du}{\sqrt{u}} \\
&= \int \frac{u^{-\frac{1}{2}}}{2} du \\
&= 2\frac{u^{\frac{1}{2}}}{2} + c \\
&= u^{\frac{1}{2}} + c \\
&= (1+x^2)^{\frac{1}{2}} + c \\
&= \sqrt{1+x^2} + c
\end{aligned}$$

A better way to compute this is what we call "advanced guessing". Once you've done enough of these problems that you know what's going to happen, you can look at the $\sqrt{1+x^2}$ in the denominator and guess that the answer will involve $(1+x^2)^{1/2}$. Once you've made a guess, differentiate it and see if it works!

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

As you can see, using this method we quickly confirm that:

$$\int \frac{xdx}{\sqrt{1+x^2}} = (1+x^2)^{1/2} + c.$$

This method is highly recommended, but it takes some getting used to.

Example: $\displaystyle\int e^{6x}dx$

We know that the derivative of e^x is e^x, so we guess e^{6x}. Then we check our guess using the chain rule:

$$\frac{d}{dx}(e)^{6x} = e^{6x}(6) = 6e^{6x}$$

This has a multiple of 6 that's not in the integral we're trying to compute, so we should divide our guess by 6 to get the correct answer:

$$\int e^{6x}dx = \frac{1}{6}e^{6x} + c.$$

We could also have used the substitution $u = 6x$. It would have worked, but it would have taken much longer.

More Examples of Integration

Example: $\int xe^{-x^2}\,dx$

For this we guess e^{-x^2}, hoping that the chain rule will somehow provide the missing factor of x in the integral. As usual, we take the derivative to check:

$$\frac{d}{dx}e^{-x^2} = (e^{-x^2})(-2x) = -2xe^{-x^2}$$

We're off by a factor of -2, so we divide our original guess by this constant to reach the conclusion that:

$$\int xe^{-x^2}\,dx = -\frac{1}{2}e^{-x^2} + c$$

Caution: If you solve integrals by guessing and don't check your answer by taking a derivative you're likely to make mistakes.

Example: $\int \sin x \cos x\,dx$

What's a good guess?

Student: $\sin^2 x$

Let's check it!

$$\frac{d}{dx}\sin^2 x = 2\sin x\cos x.$$

So:

$$\int \sin x\cos x\,dx = \frac{1}{2}\sin^2 x + c$$

An equally acceptable answer is:

$$\int \sin x\cos x\,dx = -\frac{1}{2}\cos^2 x + c$$

This seems like a contradiction; let's check our answer:

$$\frac{d}{dx}\cos^2 x = (2\cos x)(-\sin x) = -2\sin x\cos x$$

Both answers are correct! But we just proved that integrals are unique up to a constant. What's going on?

It turns out that the difference between the two answers *is* a constant:

$$\frac{1}{2}\sin^2 x - (-\frac{1}{2}\cos^2 x) = \frac{1}{2}(\sin^2 x + \cos^2 x) = \frac{1}{2}$$

So,

$$\frac{1}{2}\sin^2 x - \frac{1}{2} = \frac{1}{2}(\sin^2 x - 1) = \frac{1}{2}(-\cos^2 x) = -\frac{1}{2}\cos^2 x$$

The two answers are, in fact, equivalent. The constant c is shifted by $\frac{1}{2}$ from one answer to the other.

Example: $\displaystyle\int \frac{dx}{x \ln x}$

We will assume $x > 0$ so that $\ln x$ is defined. We don't quickly come up with a good guess, so we use the method of substitution (which is the only other method we know). The ugliest part of the integral is the natural log, so we choose:

$$u = \ln x.$$

One advantage of this choice is that taking the differential of $\ln x$ makes it simpler: $du = \dfrac{1}{x}\,dx$. Substitute these into the integral to get:

$$\begin{aligned}
\int \frac{dx}{x \ln x} &= \int \underbrace{\frac{1}{\ln x}}_{\frac{1}{u}} \underbrace{\frac{dx}{x}}_{du} \\
&= \int \frac{1}{u} du \\
&= \ln |u| + c \\
&= \ln |\ln(x)| + c
\end{aligned}$$

For this example, the method of substitution is better than guessing.

Introduction to Ordinary Differential Equations

MIT has an entire course on differential equations called 18.03. However, there is a technique using differentials that fits in well with what we've been doing with integration. We'll discuss that here.

The simplest type of differential equation looks like: $\dfrac{dy}{dx} = f(x)$. The solution to this equation is the antiderivative (integral) $y = \displaystyle\int f(x)\,dx$. We're going to assume for now that we can always solve this problem.

At the moment we know only the method of substitution (which includes "advanced guessing") for solving integration problems.

Example: $\left(\dfrac{d}{dx} + x\right) y = 0 \quad \left(\text{or} \quad \dfrac{dy}{dx} + xy = 0\right)$

This is our first interesting example of a differential equation. The operation $\left(\dfrac{d}{dx} + x\right)$ is known in quantum mechanics as the *annihilation operator*. This is also the equation that governs the ground state of the harmonic oscillator, and it is relatively simple to solve.

The first step in solving it is to rewrite the equation by isolating $\frac{dy}{dx}$:

$$\begin{aligned} \left(\frac{d}{dx} + x\right) y &= 0 \\ \frac{dy}{dx} + xy &= 0 \\ \frac{dy}{dx} &= -xy \end{aligned}$$

The big difference between this example and the antiderivatives we've been studying is that here the rate of change depends upon both x and y. We don't yet have any strategies for solving this sort of equation.

However, it turns out that we can use differentials and Leibnitz' notation to solve this. The key step is to *separate variables* so that all the terms involving y are on one side of the equation and all terms involving x are on the other.

$$\frac{dy}{y} = -x\,dx$$

Because the problem is now set up in terms of differentials, as opposed to ratios of differentials (rates of change). Because of this, we can take advantage of Leibnitz' notation and integrate both sides of the equation. Notice that on the left the differential variable is y and on the right it is x.

$$\begin{aligned} \int \frac{dy}{y} &= -\int x\,dx \\ \ln y + c_1 &= -\frac{x^2}{2} + c_2 \quad (\text{assume } y > 0) \end{aligned}$$

$$\begin{aligned}
\ln y &= -\frac{x^2}{2} + c \quad (\text{we only need one constant } c = c_2 - c_1) \\
e^{\ln y} &= e^{c+-x^2/2} \\
y &= e^c e^{-x^2/2} \\
y &= Ae^{-x^2/2} \quad (A = e^c)
\end{aligned}$$

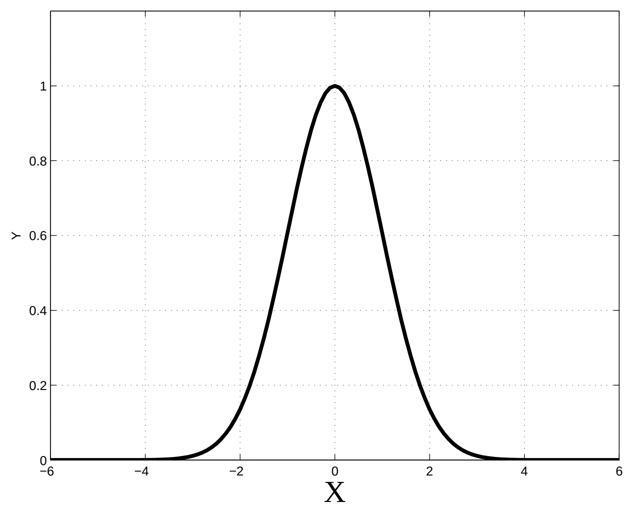


Figure 1: Graph of $y = e^{-\frac{x^2}{2}}$.

It turns out that the solution is $y = ae^{-x^2/2}$ for any constant multiple a, despite the fact that $e^c \neq 0$. We can check this solution using differentiation:

$$\begin{aligned}
y &= ae^{-x^2/2} \\
\frac{dy}{dx} &= \frac{d}{dx} ae^{-x^2/2} \\
&= ae^{-x^2/2} \cdot -2x/2 \quad (\text{chain rule}) \\
&= ae^{-x^2/2} \cdot -x \\
&= y \cdot -x \\
\frac{dy}{dx} &= -xy
\end{aligned}$$

This matches the equation for $\frac{dy}{dx}$ from our first step, so $y = ae^{-x^2/2}$ is a solution to our differential equation. We didn't make any assumptions about a in this calculation, so this is a solution no matter what value a has; $a = 0$ is possible along with all $a \neq 0$, depending on the initial conditions. For instance, if $y(0) = 1$, then $y = e^{-x^2/2}$. If $y(0) = a$, then $y = ae^{-x^2/2}$ (See Fig. 1).

This function is known as the normal distribution, which you may recognize from probability. In quantum mechanics it helps describe where a particle is.

The aim of differential equations is to solve them. Just as with algebraic equations. Usually, differential equations are telling you something about the balance between an acceleration and a velocity; for example, if you're doing calculations involving air resistance. Sometimes in applied problems, formulating a differential equation to describe a situation is very important. In order to see that you chose the right formulation you must confirm that your solution fits what actually happens in the real world.

Question: How can $y = ae^{-x^2/2}$ be our final solution when we don't know what a is?

Answer: We call $y = ae^{-x^2/2}$ the *general solution*; in other words, the whole family of solutions we get by choosing different values for a is the answer to the question.

Frequently we will be given more information than just $\left(\frac{d}{dx} + x\right) y = 0$. For example, we might know that when x is 0, y is 3. Given that extra piece of information, we can nail down exactly which function is the solution:

$$\begin{aligned} y &= ae^{-x^2/2} \\ 3 &= ae^{-(0^2)/2} \\ 3 &= ae^0 \\ 3 &= a \end{aligned}$$

And so the final solution would be $y = 3e^{-x^2/2}$.

If we don't have the information you need to restrict our answer to a single function then the solution is not one function, it's a family of functions described by the different possible values of some parameter like a.

Question: Can you solve for x instead of y?

Answer: Sure! You would get the inverse function of the function that we're officially looking for but yes, it's legal. Sometimes we'll have to make do with just an implicit formula and sometimes we're stuck with x is a function of y. The way in which the solution is specified can be complicated; as you'll soon see, it's not necessarily the best thing to think y as a function of x.

Separation of Variables

We'll now look at the technique we used to solve the previous problem and discuss what other sorts of problems that method is useful for.

In general, the method of separation of variables applies to differential equations that can be written as:

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

In our previous example, $f(x) = -x$ and $g(y) = y$.

The key step in the method is the separation of variables. It is possible because Leibnitz designed his notation to make this work nicely, allowing us to treat differential calculations as if they were ordinary arithmetic.

$$\begin{aligned} \frac{dy}{dx} &= f(x)g(y) \\ \frac{dy}{g(y)} &= f(x)\,dx \quad \text{which we can write as} \\ h(y)\,dy &= f(x)\,dx \ \text{ where } \ h(y) = \frac{1}{g(y)}. \end{aligned}$$

Next we antidifferentiate both sides of the equation:

$$H(y) = \int h(y)\,dy; \quad F(x) = \int f(x)\,dx$$

In our example, $H(y) = \ln|y|$ and $F(x) = \frac{-x^2}{2}$.

These antiderivatives are equal, so we get:

$$H(y) = F(x) + c \quad \text{(Again, we only need one constant } c.)$$

This is what we call an *implicit equation*; in our example we had $\ln y = -x^2/2 + c$. It doesn't tell us exactly what y is but it does describe y implicitly. In order to solve for y explicitly we need the inverse function H^{-1}; in our example this inverse was the exponential function and the explicit equation was $y = Ae^{-x^2/2}$.

In practice, it's often easy to find the implicit equation and quite messy to perform the inverse operation. In that case we might leave the solution in implicit form.

Remark 1: In the example, we could have written $\ln|y| = -x^2/2 + c$ for $y \neq 0$. Then we would have gotten $|y| = Ae^{-x^2/2}$ or $y = \pm Ae^{-x^2/2}$, which is almost exactly the answer we did get, if $a = \pm A$ and $A > 0$.

Professor Jerison didn't bother with this because it makes the calculation more complicated. Once he had solved the problem for $y > 0$, he knew from previous experience what the answer would be and he could skip directly to that and check his work. (The exponential function comes up all the time, so you too will want to be completely comfortable dealing with it.)

This still leaves out the case $y = 0$. This is an extremely boring solution, but it is still a solution to this problem. You can verify that $y = 0$ (i.e. $a = 0$) is a solution to this problem by plugging 0 in for y in the equation $\left(\frac{d}{dx} + x\right) y = 0$ or $\frac{dy}{dx} = -xy$.

It's not so surprising that we missed that solution, because in the process of separating variables we divided by y. If you divide by something you may have problems when that thing equals zero, or miss that solution. To avoid these problems, take note of when you divide by something that might be zero and double check that case after you've finished your calculations.

Remark 2: We had:

$$\int h(y)\,dy = \int f(x)\,dx$$

which evaluated to:

$$H(y) = F(x) + c.$$

We could have written $H(y) + c_1 = F(x) + c_2$, but this is equivalent to $H(y) = F(x) + c_2 - c_1 = F(x) + c$. To save time and writing, we write down only one arbitrary constant when integrating both sides of an equation.

Remark 3: In our example, the additive constant c turned into a multiplicative constant A when we calculated $e^{-x^2/2+c}$. In general there will always be a free constant in the solution to a differential equation, but that constant will not always be additive.

Example: $\dfrac{dy}{dx} = f(x)$

We'll solve our first, very simple, example using the method of separation of variables. We start by multiplying both sides by dx:

$$dy = f(x)\,dx.$$

Then integrate both sides:

$$\int dy = \int f(x)\,dx$$

The antiderivative of dy is just y, so we get:

$$y = \int f(x)\,dx,$$

as we did before.

Differential Equations and Slope, Part 1

Suppose the tangent line to a curve at each point (x, y) on the curve is twice as steep as the ray from the origin to that point. Find a general equation for this curve. (See Fig. 1.)

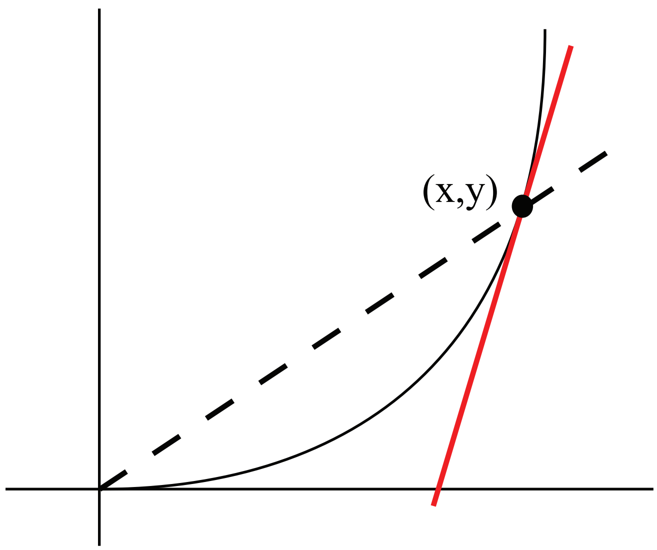


Figure 1: The slope of the tangent line (red) is twice the slope of the ray from the origin to the point (x, y).

This type of problem can be described very succinctly using differential equations. The slope of the tangent line is $\frac{dy}{dx}$. The slope of the ray from $(0, 0)$ to (x, y) is $\frac{y}{x}$. Since the slope of that ray is twice the slope of that ray, we get the differential equation:

$$\frac{dy}{dx} = 2\left(\frac{y}{x}\right).$$

We only have one method for solving differential equations; use it.

$$\begin{aligned}
\frac{dy}{dx} &= 2\frac{y}{x} \\
\frac{dy}{y} &= \frac{2\,dx}{x} \quad \text{(separate variables)} \\
\int \frac{dy}{y} &= \int \frac{2}{x}\,dx \quad \text{(integrate both sides)} \\
\ln|y| &= 2\ln|x| + c \quad \text{(antidifferentiate)} \\
e^{\ln|y|} &= e^{2\ln|x|+c} \quad \text{(apply an inverse function to isolate y)} \\
e^{\ln|y|} &= e^{c}e^{2\ln|x|} \quad \text{(exponentiate)}
\end{aligned}$$

$$
\begin{aligned}
e^{\ln|y|} &= e^c(e^{\ln|x|})^2 \\
|y| &= e^c x^2 \quad (e^{2\ln|x|} = x^2\ )
\end{aligned}
$$

There is an absolute value in this solution. When $y > 0$ we get $y = e^c x^2$. When $y < 0$ we get $y = -e^c x^2$. Based on prior experience we guess that the solution will be $y = ax^2$, where $a = \pm e^c$ or $a = 0$.

Because we divided by y in our calculations our solution doesn't include the case in which $a = 0$ and $y = 0x^2$. Graph the equation $y = 0$ and confirm that at each point on the graph the slope of the tangent line is twice the slope of the ray joining that point to the origin; this confirms that $y = 0x^2$ is a solution.

We conclude that the general solution to the problem is:

$$y = ax^2$$

where a could be positive, negative or zero. Some possible solutions include:

$$
\begin{aligned}
y &= x^2 \quad (a = 1) \\
y &= 2x^2 \quad (a = 2) \\
y &= -x^2 \quad (a = -1) \\
y &= 0x^2 = 0 \quad (a = 0) \\
y &= -2y^2 \quad (a = -2) \\
y &= 100x^2 \quad (a = 100)
\end{aligned}
$$

Some representatives of this family of curves are shown in black in Fig. 2.

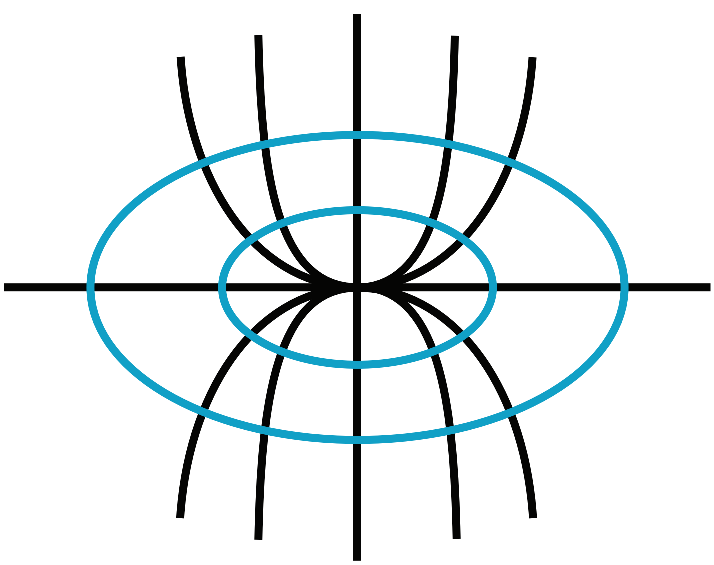


Figure 2: Parabolic curves, shown in black.

If we want to check our work, we can do so by taking the derivative:

$$\begin{aligned} y &= ax^2 \\ \frac{dy}{dx} &= 2ax \end{aligned}$$

Since $2ax = \frac{2ax^2}{x}$, we have $\frac{dy}{dx} = \frac{2y}{x}$. This works for $a > 0$, $a < 0$ and $a = 0$, so this solution is valid for all those values of a.

Warning: Notice that in the equation $\frac{dy}{dx} = \frac{2y}{x}$, $\frac{2y}{x}$ is undefined at $x = 0$. As you can see from Fig. 2, knowing the value of the function and its derivative at $x = 0$ doesn't tell us how the function will behave elsewhere. This is bad — for one thing, it contradicts our understanding of linear approximation.

What goes wrong is that the rate of change is not specified when $x = 0$. If you think carefully about what this function is doing, it could follow one branch when $x < 0$ and a completely different branch when $x > 0$. That's a very subtle point; you won't be asked to discuss this problem in your homework, but you should be aware that when x is equal to zero there's a problem for this differential equation.

Differential Equations and Slope, Part 2

Find the curves that are perpendicular to the parabolas $y = ax^2$ from the previous example.

We get a new differential equation from the one in the last example by using the fact that if a line has slope m, a line perpendicular to it will have slope $-\frac{1}{m}$. So:

$$\begin{aligned} \text{slope of curve} &= \frac{dy}{dx} \\ &= -\frac{1}{\text{slope of parabola}} \\ &= -\frac{1}{\frac{2y}{x}} \\ \frac{dy}{dx} &= \frac{-x}{2y} \end{aligned}$$

Separate variables:

$$2y\,dy = -x\,dx$$

Take the antiderivative:

$$\begin{aligned} \int 2y\,dy &= \int -x\,dx \\ y^2 &= -\frac{x^2}{2} + c \end{aligned}$$

So the general solution to this differential equation is:

$$y^2 + \frac{x^2}{2} = c.$$

This describes a family of ellipses. The y-semi-minor axis of these ellipses has length $\sqrt{c}$ and the x-semi-major axis has length $\sqrt{2c}$; the ratio of the x-semi-major axis to the y-semi-minor axis is $\sqrt{2}$ (see Fig. 1).

Unlike last time, this solution only works when $c > 0$. For some problems your constant parameter can be any real value; for some it can't.

Separation of variables leads to implicit formulas for y, but in this case you can solve for y.

$$y = \pm\sqrt{c - \frac{x^2}{2}}$$

Writing the solution in this form brings an important point to our attention — the equation of an ellipse does not describe a function! The explicit solution gives you functions that describe the top and bottom halves of the ellipses

The explicit solution also suggests that there's a problem when $y = 0$ and $x = \pm\sqrt{2c}$. Here the ellipse has a vertical tangent line; also the explicit solution isn't defined for $|x| > \sqrt{2c}$. This makes sense when we consider the fact that $\frac{dy}{dx} = \frac{-x}{2y}$. When $y = 0$ the slope of the tangent line to the curve should be infinite.

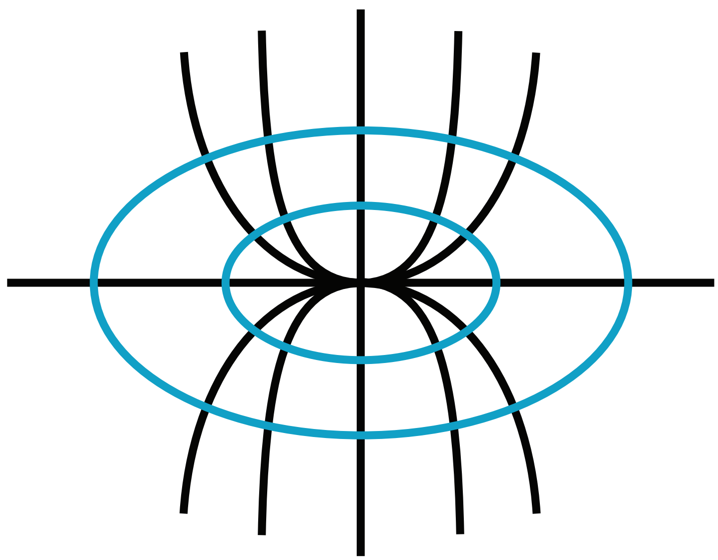


Figure 1: The curves perpendicular to the parabolas are ellipses.

Review for Test 2

Exam 2 is usually the hardest test of the course. The topics covered will be:

1. Linear and/or quadratic approximations
2. Sketches of $y = f(x)$
3. Max/min problems.
4. Related rates.
5. Antiderivatives. Separation of variables.
6. Mean value theorem.

You should also work some practice tests to familiarize yourself with the different types of questions on the exam before you take it.

In conclusion, the main theme of this unit on applications of derivatives is that information about a function's derivative gives us information about the function itself.

Introduction to Definite Integrals

As usual, we'll introduce this topic from a geometric point of view. Geometrically, definite integrals are used to find the area under a curve. Alternately, you can think of them as a "cumulative sum" — we'll see this viewpoint later.

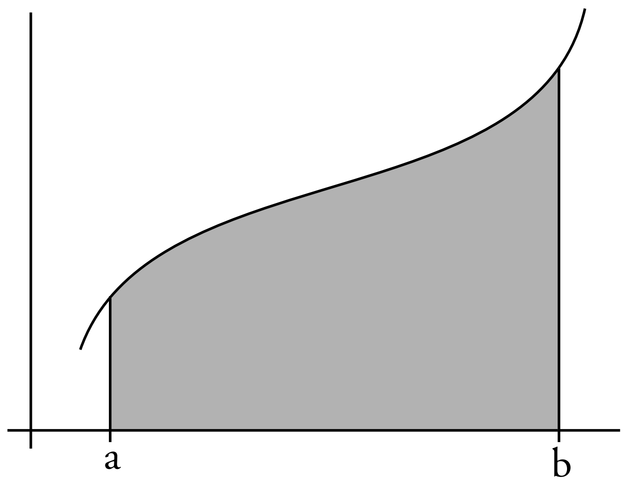


Figure 1: Area under a curve

Figure 1 illustrates what we mean by "area under a curve". The area starts at the left endpoint $x = a$ and ends at the right endpoint $x = b$. The "top" is the graph of $f(x)$ and the "bottom" is the x-axis. The notation we use to describe this in calculus is the *definite integral*

$$\int_a^b f(x)dx.$$

The difference between a definite integral and an indefinite integral (or antiderivative) is that a definite integral has specified start and end points.

Definition of the Definite Integral

The definite integral $\int_a^b f(x)dx$ describes the area "under" the graph of $f(x)$ on the interval $a < x < b$.

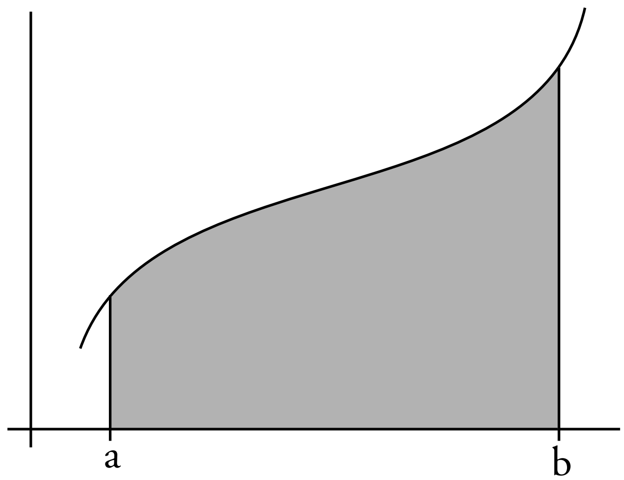


Figure 1: Area under a curve

Abstractly, the way we compute this area is to divide it up into rectangles then take a limit. The three steps in this process are:

1. Divide the region into "rectangles"

2. Add up areas of rectangles

3. Take the limit as the rectangles become infinitesimally thin

Figure 2 shows the area under a curve divided into rectangles. Notice that since the rectangles aren't curved they do not exactly overlap the area. Adding up the areas of the rectangles doesn't give you exactly the area under the curve, but the two areas are pretty close together.

The key idea is that as the rectangles get thinner, the difference between the area covered by the rectangles and the area under the curve will get smaller. In the limit, the area covered by the rectangles will exactly equal the area under the curve.

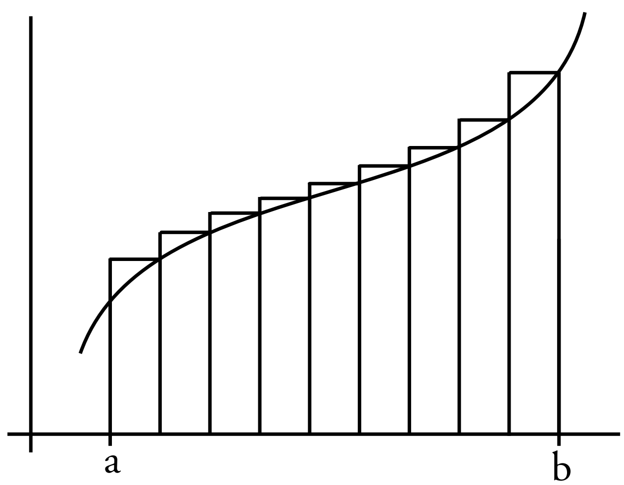


Figure 2: Area under a curve divided into rectangles

Example: $f(x) = x^2$

Professor Jerison only does one simple example of computing the definite integral as a limit because computing integrals this way involves a lot of hard work.

In this example we'll use the first interesting curve, $f(x) = x^2$, with starting point $a = 0$. In order to see what the pattern is, we'll allow b to be arbitrary — we can replace it with a value after the calculation if we like.

We start by graphing $f(x)$ and identifying the region whose area we are computing.

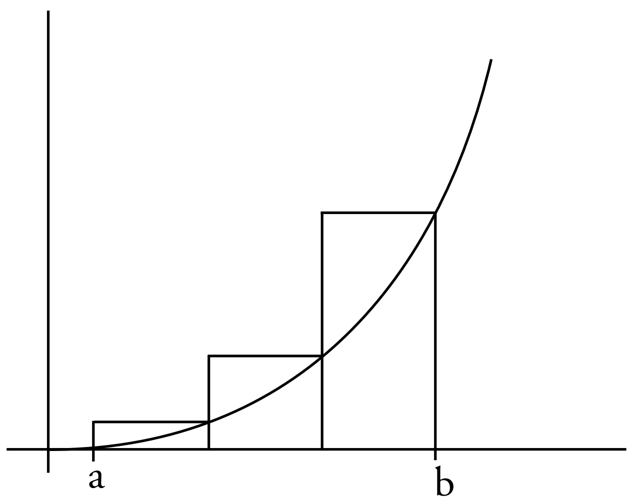


Figure 1: $\int_a^b f(x)dx$ approximated by three rectangles

Next, we subdivide the interval from 0 to b into n pieces; in this class the pieces will always have equal length. An example with $n = 3$ is show in Figure 1. Each subdivision forms the base of a rectangle. The tops of the rectangles touch the graph; Professor Jerison has chosen to have the intersection of the rectangle and the graph in the upper right corner of the rectangle. Because of this, the sum of the areas of the rectangles will be slightly larger than the area under the curve.

Next, we find the areas of the rectangles. The nice thing about rectangles is that it's easy to compute their areas — just multiply the length of the base by the height. Since each of the n rectangle bases is the same size, the base of each rectangle has length $\dfrac{b}{n}$. The height of each rectangle will be given by $f(x_i) = x_i^2$, where x_i is the right endpoint of the base.

The first rectangle has base $\left[0, \frac{b}{n}\right]$, so its height is $\left(\frac{b}{n}\right)^2$ and its area is $\left(\frac{b}{n}\right)^3$. Information about other rectangles appears in the table below.

x (right endpoint of base)	$f(x)$ (height of rectangle)
$\frac{b}{n}$	$\left(\frac{b}{n}\right)^2$
$\frac{2b}{n}$	$\left(\frac{2b}{n}\right)^2$
$\frac{3b}{n}$	$\left(\frac{3b}{n}\right)^2$
$\vdots$	$\vdots$
$\frac{nb}{n}$	b^2

Now we can easily find the areas of the rectangles; then we add them up to get an approximation of the area under the curve:

$$\underbrace{\left(\frac{b}{n}\right)}_{\text{base}}\underbrace{\left(\frac{b}{n}\right)^2}_{\text{height}} + \left(\frac{b}{n}\right)\left(\frac{2b}{n}\right)^2 + \left(\frac{b}{n}\right)\left(\frac{3b}{n}\right)^2 + \cdots + \left(\frac{b}{n}\right)\left(\frac{nb}{n}\right)^2$$

This is complicated. We'll start by simplifying it and work toward evaluating the limit as n goes to infinity (i.e. when the base of the rectangle is infinitesimal.) It turns out that it's easier to calculate the limit than to calculate the sum of the areas of the rectangles; that's why we study calculus.

First, we can factor out $\left(\frac{b}{n}\right)^3$:

$$\begin{aligned}\left(\frac{b}{n}\right)\left(\frac{b}{n}\right)^2 + \left(\frac{b}{n}\right)\left(\frac{2b}{n}\right)^2 + \left(\frac{b}{n}\right)\left(\frac{3b}{n}\right)^2 + \cdots + \left(\frac{b}{n}\right)\left(\frac{nb}{n}\right)^2 &= \\ \left(\frac{b}{n}\right)\frac{b^2}{n^2} + \left(\frac{b}{n}\right)\frac{2^2b^2}{n^2} + \left(\frac{b}{n}\right)\frac{3^2b^2}{n^2} + \cdots + \left(\frac{b}{n}\right)\frac{n^2b^2}{n^2} &= \\ \frac{b^3}{n^3}(1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2)\end{aligned}$$

We want to take the limit as n goes to infinity. What makes this difficult is the sum $1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2$. We're going to use a geometric "trick" to draw a picture representing this quantity.

Geometric Calculation of Sum of Squares:

Imagine you're building a pyramid. The base of the pyramid is an n by n square of n^2 cubes. The next layer is $n-1$ by $n-1$, and has $(n-1)^2$ blocks in it. The top view of the pyramid shows a series of concentric squares. In profile, the pyramid looks like a triangle formed out of rectangles with height 1 and length n, $(n-1)$, etc. The left and right sides of that triangle have slopes 2 and -2.

The volume of this pyramid is $n^2 + (n-1)^2 + ... + 3^2 + 2^2 + 1^2$, which equals the difficult sum from our definite integral. This volue is slightly larger than that of the ordinary pyramid with base n and height n which is the largest ordinary pyramid contained entirely inside our stair-step pyramid. We know that the volume of that inside pyramid is $\frac{1}{3} \cdot \text{base} \cdot \text{height}$, or $\frac{1}{3}n^2 \cdot n$. So

$$\frac{1}{3}n^3 < 1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2.$$

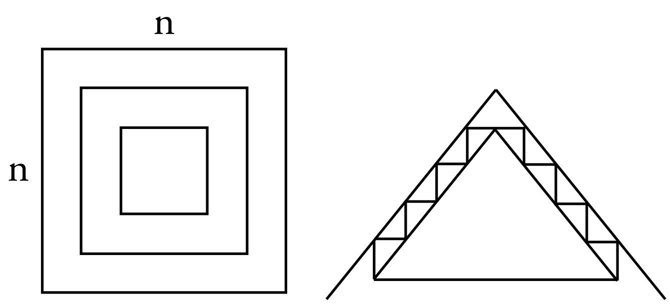


Figure 2: Top and side views of a stair-step pyramid.

We've compared our difficult sum, which equaled the volume of a stair-step pyramid, to the volume $\frac{1}{3}n^3$ of the biggest ordinary pyramid that fit inside it. This is a step toward replacing the difficult sum by something much simpler.

Next we compare the volume of the stair-step pyramid to the volume of the smallest ordinary pyramid that can contain it. That pyramid has base $n+1$ and height $n+1$, and so has volume $\frac{1}{3}(n+1)^3$. Hence:

$$\frac{1}{3}n^3 < 1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2 < \frac{1}{3}(n+1)^3.$$

We now divide everything by n^3 to get:

$$\frac{1}{3} < \frac{1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2}{n^3} < \frac{1}{3}\frac{(n+1)^3}{n^3} = \frac{1}{3}\left(1 + \frac{1}{n}\right)^3.$$

If we let n go to infinity, we find that the left and right sides of this inequality approach $\frac{1}{3}$ and so the center expression

$$\frac{1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2}{n^3} \longrightarrow \frac{1}{3} \text{ as } n \longrightarrow \infty.$$

The sum of the areas of the rectangles under the graph of x^2 was

$$\frac{b^3}{n^3}(1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2).$$

As n approaches infinity, this area approaches $\frac{b^3}{3}$. So the total area between the graph of $f(x) = x^2$ and the interval $[0, b]$ is:

$$\int_0^b x^2 dx = \frac{1}{3}b^3.$$

Question: Why did we leave the $\left(\frac{b}{n}\right)^3$ out for this step?

Answer: Part of the answer is that we know what we're heading for. We understand the quantity $\left(\frac{b}{n}\right)^3$. However, the difficult sum $(1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2)$ is growing larger and larger in a way we don't entirely understand. So we separate out the difficult sum and concentrate on that. When we do, we discover that it's very, very similar to n^3, and is even more similar to $\frac{1}{3}n^3$. Once we understand that we can use it in the original equation, cancel the n^3's, and get our result.

This is what you always do if you analyze these kinds of sum; you factor out whatever you understand and end up with a sum like this. You should expect this to happen every time you're faced with such a sum.

In summary, the steps we followed to find the area under the curve were:

1. Graph the function

2. Subdivide into n intervals of length $\frac{b-a}{n}$

3. Compute the heights $f\left(\frac{i(b-a)}{n}\right)$ of the rectangles

4. Compute the areas of the rectangles

5. Sum the areas of the rectangles

6. Find the limit of this sum as n goes to infinity.

Summation Notation

You'll have noticed working with sums like $1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2$ is extremely cumbersome; it's really too large for us to deal with. Mathematicians have a shorthand for calculations like this which doesn't make the arithmetic any easier, but does make it easier to write down these sums.

The general notation is:

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n.$$

The summation symbol Σ is a capital sigma. So, for instance,

$$\frac{1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2}{n^3} = \frac{1}{n^3} \sum_{i=1}^{n} i^2.$$

We just showed that:

$$\lim_{n\to\infty} \frac{1}{n^3} \sum_{i=1}^{n} i^2 = \frac{1}{3}.$$

When using the summation notation, we'll have a formula describing each summand a_i in terms of i; for example, $a_i = i^2$. The expression $\sum_{i=1}^{n} a_i$ is just an abbreviation for the sum of the terms a_i.

Another difficult sum we encountered was:

$$\left(\frac{b}{n}\right)\left(\frac{b}{n}\right)^2 + \left(\frac{b}{n}\right)\left(\frac{2b}{n}\right)^2 + \left(\frac{b}{n}\right)\left(\frac{3b}{n}\right)^2 + \cdots + \left(\frac{b}{n}\right)\left(\frac{nb}{n}\right)^2$$

Using summation notation, we can rewrite this as:

$$\sum_{i-1}^{n} \left(\frac{b}{n}\right)\left(\frac{ib}{n}\right)^2.$$

We factored $\left(\frac{b}{n}\right)^3$ out of this sum earlier; we can also do this using our new notation:

$$\sum_{i-1}^{n} \left(\frac{b}{n}\right)\left(\frac{ib}{n}\right)^2 = \frac{b^3}{n^3} \sum_{i=1}^{n} i^2.$$

These notations just make our notes a little bit more compact. The concepts are still the same and the mess is still there hiding under the rug, but the notation at least fits on the page.

Easy Definite Integrals

We'll do two more (much easier) examples so that we can see the pattern in these calculations.

Example: $f(x) = x$

Compute $\int_0^b f(x)\,dx$.

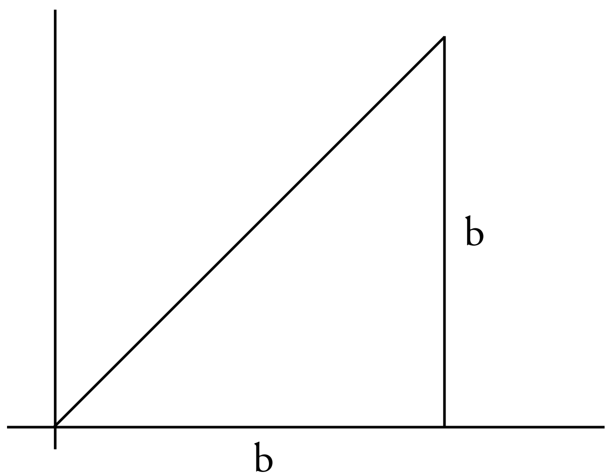


Figure 1: The area under the curve is $\int_0^b x dx$

Looking at Figure 1 we see that the area under the curve is just a triangle with area

$$\frac{1}{2}\underbrace{b}_{\text{base}}\cdot\underbrace{h}_{\text{ht}} = \frac{1}{2}b^2.$$

We conclude that $\int_0^b x\,dx = \frac{1}{2}b^2$ *without* doing any elaborate summing, because we happen to know this area.

Example: $f(x) = 1$

This is by far the most important example, but by the time you get to 18.02 and multivariable calculus you will forget this calculation.

The graph of the function is just a horizontal line. The area under that line between 0 and b is the area of a rectangle with length b and height 1. In other words,

$$\int_0^b 1\,dx = b.$$

Summary of Examples

Now let's look at our results, comparing the function $f(x)$ to the area $\int_0^b f(x)\,dx$ under the graph of f between 0 and b.

$f(x)$	$\int_0^b f(x)\,dx$
x^2	$b^3/3$
$x = x^1$	$b^2/2$
$1 = x^0$	$b = b^1/1$

It looks as if a good guess for $\int_0^b x^3\,dx$ should be $b^4/4$, and in fact this guess is correct.

Historically, Archimedes figured out the area under a parabola in the third century B.C. He used a much more complicated method than is described here, and his method was so brilliant that it may have set back mathematics by 2,000 years. It was so difficult that people couldn't see this pattern, and couldn't see that these kinds of calculations can be easy. They couldn't get to the cubic, and even when they did they were struggling with everything else. It wasn't until calculus fit everything together that people were able to make serious progress on calculating these areas.

We now have easy methods for computing these volumes; we will not have to labor to build pyramids to calculate all of these quantities. We will be able to do it so easily that it will happen as fast as you differentiate.

Riemann Sums

We haven't yet finished with approximating the area under a curve using sums of areas of rectangles, but we won't use any more elaborate geometric arguments to compute those sums.

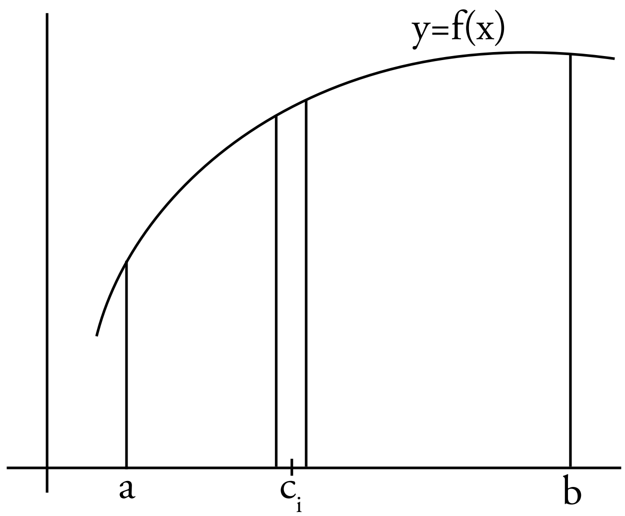


Figure 1: Area under a curve

The general procedure for computing the definite integral $\int_a^b f(x)\,dx$ is:

- Divide $[a, b]$ into n equal pieces of length $\Delta x = \dfrac{b-a}{n}$.
- Pick *any* value c_i in the i^{th} interval and use $f(c_i)$ as the height of the rectangle.
- Sum the areas of the rectangles:

$$f(c_1)\Delta x + f(c_2)\Delta x + \cdots + f(c_n)\Delta x = \sum_{i=1}^{n} f(c_i)\Delta x$$

The sum $\sum_{i=1}^{n} f(c_i)\Delta x$ is called a *Riemann Sum.*

This notation is supposed to be reminiscent of Leibnitz' notation. In the limit as n goes to infinity, this sum approaches the value of the definite integral:

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\Delta x = \int_a^b f(x)\,dx$$

Which is the area under the curve $y = f(x)$ above $[a, b]$.

Example: Cumulative Debt

In this example we see an integral that represents a cumulative sum, rather than an area.

Let t = time in years
and $f(t)$ = dollars/year; $f(t)$ is a borrowing rate.

Notice the units in this problem; they are one of the reasons we include a differential like dx in all of our integrals. This notation to be consistent with units, helps in changing variables, and allows us to develop meaningful formulas which are consistent across the board.

Suppose you're borrowing money every day; then $\Delta t = \frac{1}{365}$ years. In terms of years, this is a nearly infinitesimal interval of time. Your borrowing rate varies over the year; you borrow more when you need more, and less when you need less. How much do you borrow?

On Day 45, which is at $t = 45/365$, you borrowed $f\left(\frac{45}{365}\right)\Delta t = f\left(\frac{45}{365}\right)\frac{1}{365}$. Here $f(t)$ is measured in dollars per year and Δt is measured in years, so $f\left(\frac{45}{365}\right)\frac{1}{365}$ is a number of dollars; in fact it's the amount that you actually borrow on the 45^{th} day.

The total amount borrowed in an entire year is:

$$\sum_{i}^{365} f\left(\frac{i}{365}\right)\Delta t.$$

This is a messy sum, but your bank knows how to keep track of it. However, when we're modeling trading strategies of course and trying to cleverly optimize how much you borrow, how much you spend, and how much you invest you will want to replace it by $\int_0^1 f(t)\,dt$. If $\Delta t = \frac{1}{365}$, this is probably a good enough approximation.

But we're not done yet; it's equally important to model how much you owe, in addition to how much you borrowed. Since we assumed that we could approximate the Riemann sum by an integral, we'll also assume that the interest on our debt is compounded continuously. If we start with a debt of P, then after time t you owe Pe^{rt}, where r is the interest rate. Let's assume you're borrowing at a 5% interest rate; then $r = 0.05$ per year.

Over the course of the year you borrowed the amounts $f\left(\frac{i}{365}\right)\Delta t$. When you borrow this amount, the amount of time left in the year is $T = 1 - \frac{i}{365}$, which is the amount of time this incremental debt will accumulate interest. So a debt of $f\left(\frac{i}{365}\right)\Delta t$ on day i increases to a debt of:

$$\left(f\left(\frac{i}{365}\right)\Delta t\right)e^{r\left(1-\frac{i}{365}\right)}$$

at the end of the year. This is the term you sum to get your total debt at the end of the year:

$$\sum_{1}^{365}\left(f\left(\frac{i}{365}\right)\Delta t\right)e^{r\left(1-\frac{i}{365}\right)}\longrightarrow\int_0^1 e^{r(1-t)}f(t)\,dt.$$

If you're trying to decide a borrowing strategy, you're faced with integrals of this type.

Example: Cumulative Debt

In this example we see an integral that represents a cumulative sum, rather than an area.

Let t = time in years
and $f(t)$ = dollars/year; $f(t)$ is a borrowing rate.

Notice the units in this problem; they are one of the reasons we include a differential like dx in all of our integrals. This notation to be consistent with units, helps in changing variables, and allows us to develop meaningful formulas which are consistent across the board.

Suppose you're borrowing money every day; then $\Delta t = \frac{1}{365}$ years. In terms of years, this is a nearly infinitesimal interval of time. Your borrowing rate varies over the year; you borrow more when you need more, and less when you need less. How much do you borrow?

On Day 45, which is at $t = 45/365$, you borrowed $f\left(\frac{45}{365}\right)\Delta t = f\left(\frac{45}{365}\right)\frac{1}{365}$. Here $f(t)$ is measured in dollars per year and Δt is measured in years, so $f\left(\frac{45}{365}\right)\frac{1}{365}$ is a number of dollars; in fact it's the amount that you actually borrow on the $45^{\text{t}h}$ day.

The total amount borrowed in an entire year is:

$$\sum_{i}^{365} f\left(\frac{i}{365}\right)\Delta t.$$

This is a messy sum, but your bank knows how to keep track of it. However, when we're modeling trading strategies of course and trying to cleverly optimize how much you borrow, how much you spend, and how much you invest you will want to replace it by $\int_0^1 f(t)\,dt$. If $\Delta t = \frac{1}{365}$, this is probably a good enough approximation.

But we're not done yet; it's equally important to model how much you owe, in addition to how much you borrowed. Since we assumed that we could approximate the Riemann sum by an integral, we'll also assume that the interest on our debt is compounded continuously. If we start with a debt of P, then after time t you owe Pe^{rt}, where r is the interest rate. Let's assume you're borrowing at a 5% interest rate; then $r = 0.05$ per year.

Over the course of the year you borrowed the amounts $f\left(\frac{i}{365}\right)\Delta t$. When you borrow this amount, the amount of time left in the year is $T = 1 - \frac{i}{365}$, which is the amount of time this incremental debt will accumulate interest. So a debt of $f\left(\frac{i}{365}\right)\Delta t$ on day i increases to a debt of:

$$\left(f\left(\frac{i}{365}\right)\Delta t\right)e^{r\left(1-\frac{i}{365}\right)}$$

at the end of the year. This is the term you sum to get your total debt at the end of the year:

$$\sum_{1}^{365} \left(f\left(\frac{i}{365}\right) \Delta t \right) e^{r\left(1-\frac{i}{365}\right)} \longrightarrow \int_0^1 e^{r(1-t)} f(t)\, dt.$$

If you're trying to decide a borrowing strategy, you're faced with integrals of this type.

The Fundamental Theorem of Calculus

The fundamental theorem of calculus is probably the most important thing in this entire course. There will be two versions of it; when we need to abbreviate we'll refer to the first as FTC1 and the second as FTC2.

Theorem: If $f(x)$ is continuous and $F'(x) = f(x)$, then:

$$\int_a^b f(x)dx = F(b) - F(a).$$

This may look familiar; when we talked about antiderivatives we wrote:

$$F(x) = \int f(x)\, dx.$$

The output of an indefinite integral is a function or family of functions; the output of a definite integral is a real number. The fundamental theorem of calculus is the connection between definite and indefinite integrals.

Notation: We need not always name the antiderivative function; we can use the following abbreviation:

$$F(b) - F(a) = F(x)|_a^b = F(x)|_{x=a}^{x=b}.$$

The later form is useful when you wish to emphasize which variable you will substitute the values for.

This allows us to rewrite the fundamental theorem of calculus as:

$$\int_a^b f(x)dx = F(x)|_a^b.$$

The First Fundamental Theorem of Calculus

Our first example is the one we worked so hard on when we first introduced definite integrals:

Example: $F(x) = \frac{x^3}{3}$.

When we differentiate $F(x)$ we get $f(x) = F'(x) = x^2$. The fundamental theorem of calculus tells us that:

$$\int_a^b x^2\,dx = \int_a^b f(x)\,dx = F(b) - F(a) = \frac{b^3}{3} - \frac{a^3}{3}$$

This is more compact in the new notation. We'll use it to find the definite integral of x^2 on the interval from 0 to b, to get exactly the result we got before:

$$\int_0^b x^2\,dx = \int_0^b f(x)\,dx = F(x)|_0^b = \left.\frac{x^3}{3}\right|_0^b = \frac{b^3}{3}.$$

By using the fundamental theorem of calculus we avoid the elaborate computations, difficult sums, and evaluation of limits required by Riemann sums.

Example: Area under one "hump" of $\sin(x)$.

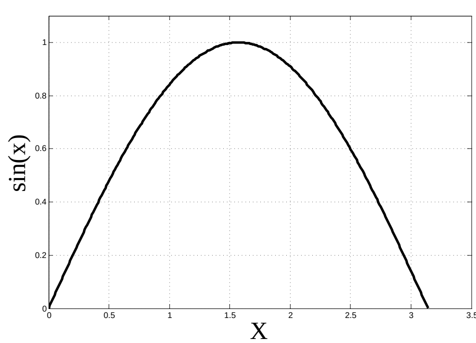


Figure 1: $\sin(x)$ for $0 < x < \pi$

The area under the curve $y = \sin x$ between 0 and π is given by the definite integral $\int_0^\pi \sin(x)\,dx$. The antiderivative of $\sin(x)$ is $-\cos(x)$, so we apply the fundamental theorem of calculus with $F(x) = -\cos(x)$ and $f(x) = \sin(x)$:

$$\int_0^\pi \sin(x)\,dx = -\cos(x)|_0^\pi .$$

Be careful with the arithmetic on the next step; it's easy to make a mistake:

$$-\cos(x)|_0^\pi = -\cos(\pi) - (-\cos(0)) = -(-1) - (-1) = 2.$$

So the area under one hump of the graph of $\sin(x)$ is simply 2 square units.

Example: $\int_0^1 x^{100}\,dx$

$$\int_0^1 x^{100}\,dx = \left.\frac{x^{101}}{101}\right|_0^1 = \frac{1}{101} - 0 = \frac{1}{101}$$

Interpretation of the Fundamental Theorem

We'll talk about a proof of the fundamental theorem later; for now let's get a more intuitive interpretation of the theorem. We'll use the example of time and distance, rather than using area again.

So, t is a time and $x(t)$ is a position at time t. The rate of change of position with respect to time is $\frac{dx}{dt} = x'(t)$, and is also known as the speed $v(t)$. We have a function $x(t)$ and it's derivative $v(t)$, so the fundamental theorem tells that:

$$\int_a^b v(t)\,dt = x(b) - x(a).$$

On the right hand side we have information about distance traveled from time $t = a$ to time $t = b$; this is what you would read on your odometer. The left hand side is what you would read on your speedometer during the trip. The fundamental theorem of calculus is telling us that if we know how fast we're going at every stage of a trip, we can figure out how far we traveled.

Let's go one step further with this interpretation to make a connection to Riemann sums. First imagine that you are extremely obsessive and while you're driving from time a to time b you check your speedometer every second. When you've read your speedometer in the i'th second, you see that you're going at speed $v(t_i)$.

How far do you go in that second? You go this speed times the time interval Δt; in this case $\Delta t = 1$ second. The distance traveled in the ith second is:

$$v(t_i)\Delta t.$$

Over the entire trip, you travel the sum of all these distances:

$$\sum_{i=1}^{n} v(t_i)\Delta t,$$

where n is some ridiculous number of seconds. The value of that sum is very close to the distance traveled recorded on your odometer because your speed doesn't change much over the course of a second. In other words,

$$\sum_{i=1}^{n} v(t_i)\Delta t \approx \int_a^b v(t)\,dt.$$

The Riemann sum, on the left, is approximately how far you traveled. The integral, on the right, is exactly how far you traveled.

The Fundamental Theorem and Negative Integrands

We said that the fundamental theorem would tell us the distance traveled if we knew the speed we were traveling at every instant. But if we make a round trip, the difference $x(b) - x(a)$ is 0. We'd like the fundamental theorem to notice whether our velocity was in the positive or negative direction and cancel the change in position when appropriate.

Goal: Extend integration to the case $f(x) < 0$.

Although we assumed that the graph of the function was above the x axis when we originally described definite integrals, it turns out that we don't have to change our notation to accommodate functions with negative outputs.

Example: $\int_0^{2\pi} \sin(x)\,dx$

Let's try evaluating the "area under" two humps of the graph of the sine function.

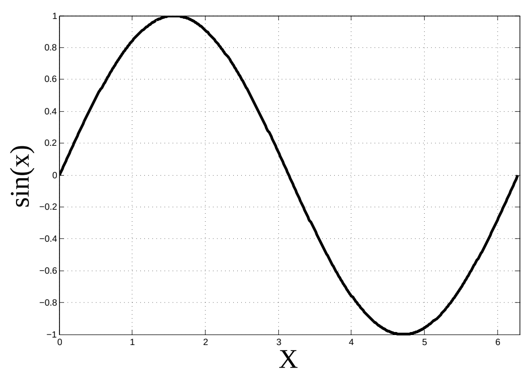


Figure 1: $\sin(x)$ for $0 < x < 2\pi$

$$\int_0^{2\pi} \sin x\,dx = (-\cos x)\Big|_0^{2\pi} = -\cos(2\pi) - (-\cos(0)) = -1 - (-1) = 0$$

If we're going to insist that the fundamental theorem of calculus must be true even when $f(x) < 0$, then the definite integral is *not* exactly the area between the curve and the x-axis. The definite integral does equal the area under the curve when the graph is above the x-axis, but when the graph is below the x-axis the value we get from the fundamental theorem is negative.

The true geometric interpretation of the definite integral is that it adds up the area above the x-axis (and below the graph of the function) and subtracts the area below the x-axis (above the graph of the function).

Question: Wouldn't you use the absolute value of the velocity?

Answer: You would use the absolute value of the velocity to compute the *total* distance traveled. Without the absolute value, the definite integral measures the *net* distance traveled.

Properties of Integrals

The symbol $\int$ originated as a stylized letter S; in French, they call integrals sums. We know from our discussion of Riemann sums that definite integrals are just limits of sums. Because of this, it's not surprising that:

1. The integral of a sum is the sum of the integrals:

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

2. We can factor out a constant multiple:

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx \quad (c \text{ constant})$$

   (don't try to factor out a non-constant function!)

3. We can combine definite integrals. If $a < b < c$ then:

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx$$

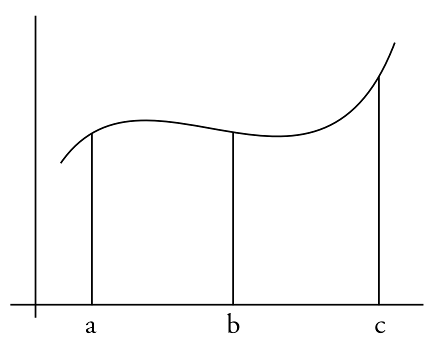


Figure 1: Combining two areas under a curve

4. $\int_a^a f(x)\,dx = 0$

5. This statement gives us some freedom in choosing limits of integration and allows us to remove the condition that $a < b < c$ from property (3):

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

   This makes sense; $F(b) - F(a) = -(F(a) - F(B))$.

6. (Estimation) If $f(x) \leq g(x)$ and $a < b$, then:

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

In other words, if I'm going more slowly than you then you go further than I do. Caution: this only works if $a < b$.

7. (Change of Variables or "Substitution") In indefinite integrals, if $u = u(x)$ then $du = u'(x)\,dx$ and $\int g(u)du = \int g(u(x))u'(x)\,dx$. To adapt this to definite integrals we need to know what happens to our limits of integration; it turns out that the answer is very simple.

$$\int_{u_1}^{u_2} g(u)du = \int_{x_1}^{x_2} g(u(x))u'(x)\,dx,$$

where $u_1 = u(x_1)$ and $u_2 = u(x_2)$. This is true if u is always increasing or always decreasing on $x_1 < x < x_2$; in other words, if u' does not change sign. (If u' does change sign you must break the integral into pieces; we'll see an example of this later.)

Example of Estimation

Here's an example in which we use the estimation property of integrals: if $f(x) \le g(x)$ and $a < b$, then $\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$.

The example is the same as one we've already seen. We'll start with an inequality and then integrate it to reach a conclusion about the antiderivatives.

We know that $e^x \ge 1$ for $x \ge 0$; this is our starting place. We integrate this expression, then follow our noses to get the result we're expecting:

$$\begin{aligned} e^x &\ge 1 \quad (x \ge 0) \\ \int_0^b e^x\,dx &\ge \int_0^b 1\,dx \quad (b \ge 0) \\ e^x\big|_0^b &\ge b \quad \text{(area of rectangle with base } b \text{ and height 1.)} \\ e^b - 1 &\ge b \\ e^b &\ge 1 + b \quad (b \ge 0) \end{aligned}$$

Notice that we can still compute the integral if $b < 0$, but in that case e^b is not greater than or equal to 1, and so we can't use the estimation property to conclude that $e^b \ge 1 + b$.

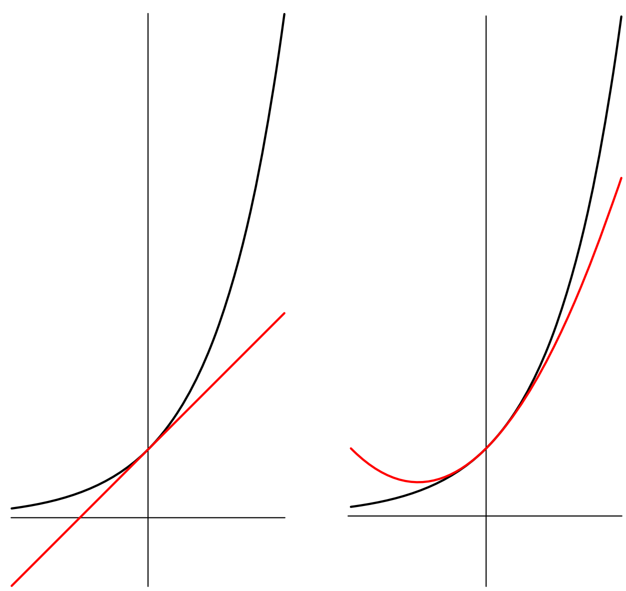


Figure 1: The graphs of e^x (black) compared to $1 + x$ and $1 + x + \frac{x^2}{2}$ (red).

Now we repeat the process starting from the conclusion:

$$\begin{aligned} e^x &\ge 1 + x \quad (x \ge 0) \\ \int_o^b e^x\,dx &\ge \int_0^b (1 + x)\,dx \quad (b \ge 0) \end{aligned}$$

$$
\begin{aligned}
e^b - 1 &\geq \left.\left(x + \frac{x^2}{2}\right)\right|_0^b \\
e^b - 1 &\geq b + \frac{b^2}{2} \\
e^b &\geq 1 + b + \frac{b^2}{2} \quad (b \geq 0)
\end{aligned}
$$

In this case, the conclusion is false if $b < 0$.

We can easily keep going with this, producing higher and higher degree interesting polynomial lower bounds for e^x. For example, if we let $b = 1$ in our final conclusion we discover that $e \geq 2\frac{1}{2}$.

Example: Change of Variables

Example: $\int_1^2 (x^3+2)^5 x^2\,dx$

Before, we would have tried to handle this integral by substitution, using $u = x^3 + 2$. We're going to do the same thing here, taking into account the limits.

First we compute $du = 3x^2$. We'll be integrating u^5, and $\frac{1}{3}\,du$ will replace $x^2\,dx$. All that's left to set up the integral is to figure out the new limits; this is one of the reasons we use dx and du — to remind ourselves which variable is involved in the integration.

Initially, x is varying between 1 and 2. So $u_1 = 1^3+2 = 3$ and $u_2 = 2^3+2 = 10$. Now we can finish the problem:

$$\begin{aligned}
\int_{x=1}^{x=2} (x^3+2)^5 x^2\,dx &= \int_{u=3}^{u=10} u^5 \frac{1}{3} u\,du \\
&= \left.\frac{u^6}{18}\right|_{u=3}^{u=10} \qquad \left(\textbf{not } \left.\tfrac{u^6}{18}\right|_1^2\right) \\
&= \frac{1}{18}(10^6 - 3^6)
\end{aligned}$$

Substitution When u' Changes Sign

We've been told that changing variables of integration only works if $u(x)$ is either always increasing or always decreasing on the interval of integration. Let's see what goes wrong by trying to calculate $\int_{-1}^{1} x^2\,dx$.

We'll try plugging in $u(x) = x^2$; then we get:

$$\begin{aligned} du &= 2x\,dx \\ dx &= \frac{1}{2x}\,du = \frac{1}{2\sqrt{u}}\,du \\ u_1 &= (-1)^2 \quad \text{and} \end{aligned}$$

$$u_2 = (-1)^2$$

Thus:

$$\int_{-1}^{1} x^2\,dx = \int_{1}^{1} u\frac{1}{2\sqrt{u}}\,du = 0.$$

But we know that $\int_{-1}^{1} x^2\,dx$ is not zero; it's the area under a parabola. Our conclusion is **not true**.

The reason for this is that $u'(x) = 2x$ is negative when $x < 0$ and positive when $x > 0$; the sign change causes us trouble. If we break the integral into two halves so that u' has a consistent sign on each half, we'll be able to compute the integral without difficulty.

We could actually have caught this early; there is a mistake in our calculation of the expression for dx. In fact, when we wrote:

$$\frac{1}{2x}\,du = \frac{1}{2\sqrt{u}}\,du$$

we should have noticed that in fact:

$$\frac{1}{2x}\,du = \frac{1}{\pm 2\sqrt{u}}\,du.$$

It's possible to use this formula to get the correct answer, but not recommended. Instead, just split your integral into intervals over which u' is always either positive or negative.

Review of the Fundamental Theorem of Calculus

Remember that the **First Fundamental Theorem of Calculus (FTC1)** said that if $F' = f$, then $\int_a^b f(x)\,dx = F(b) - F(a)$.

We used this to evaluate definite integrals; today we're going to reverse that and read the equation backward:

$$F(b) - F(a) = \int_a^b f(x)\,dx$$

and use the derivative $f = F'$ to understand the function F.

The Fundamental Theorem and the Mean Value Theorem

Our goal is to use information about F' to derive information about F. Our first example of this process will be to compare the first fundamental theorem to the Mean Value Theorem.

We'll use the notation $\Delta F = F(b) - F(a)$ and $\Delta x = b - a$. The first fundamental theorem then tells us that:

$$\Delta F = \int_a^b f(x)\,dx.$$

If we divide both sides by Δx we get:

$$\frac{\Delta F}{\Delta x} = \underbrace{\frac{1}{b-a}\int_a^b f(x)\,dx}_{Average(f)}$$

the expression on the right is the average value of the function $f(x)$ on the interval $[a, b]$.

Why is this the average of f and not of F? Consider the following Riemann sum:

$$\int_0^n f(x)\,dx \approx f(1) + f(2) + \cdots + f(n).$$

This is a cumulative sum of values of $f(x)$. The quantity:

$$\frac{\int_0^n f(x)\,dx}{n} \approx \frac{f(1) + f(2) + \cdots + f(n)}{n}$$

is an average of values of $f(x)$; in the limit, the average value of $f(x)$ on the interval $[a, b]$ is given by $\frac{1}{b-a}\int_a^b f(x)\,dx$.

We'll rewrite the first fundamental theorem one more time as:

$$\Delta F = \text{Average}(F')\Delta x.$$

In other words, the change in F is the average of the infinitesimal change times the amount of time elapsed. We can now use inequalities to compare this to the mean value theorem, which says that $\frac{F(b)-F(a)}{b-a} = F'(c)$ for some c between a and b. We can rewrite this as:

$$\Delta F = F'(c)\Delta x.$$

The value of Average(F') in the first fundamental theorem is very specific, but the $F'(c)$ from the mean value theorem is not; all we know about c is that it's somewhere between a and b.

Even if we don't know exactly what c is, we know for sure that it's less than the maximum value of F' on the interval from a to b, and that it's greater than the minimum value of F' on that interval:

$$\left(\min_{a<x<b} F'(x)\right)\Delta x \le \Delta F = F'(c)\Delta x \le \left(\max_{a<x<b} F'(x)\right)\Delta x.$$

The first fundamental theorem of calculus gives us a much more specific value — Average(F') — from which we can draw the same conclusion.

$$\left(\min_{a<x<b} F'(x)\right)\Delta x \leq \Delta F = \text{Average} F' \Delta x \leq \left(\max_{a<x<b} F'(x)\right)\Delta x.$$

The fundamental theorem of calculus is much stronger than the mean value theorem; as soon as we have integrals, we can abandon the mean value theorem. We get the same conclusion from the fundamental theorem that we got from the mean value theorem: the average is always bigger than the minimum and smaller than the maximum. Either theorem gives us the same conclusion about the change in F:

$$\left(\min_{a<x<b} F'(x)\right)\Delta x \leq \Delta F \leq \left(\max_{a<x<b} F'(x)\right)\Delta x.$$

The Mean Value Theorem and Estimation

The following problem appeared on the second exam:

Given that $F'(x) = \frac{1}{1+x}$ and $F(0) = 1$, the mean value theorem implies that $A < F(4) < B$ for which A and B?

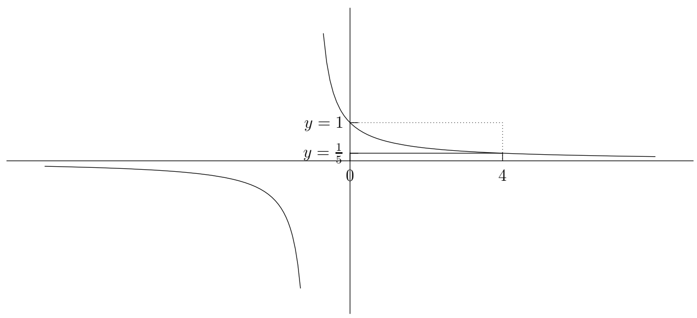


Figure 1: Graph of $F'(x) = \frac{1}{1+x}$.

To solve this, we first apply the mean value theorem in such a way that the value $F(4)$ appears, then use our knowledge of the formula for $F'(c)$ to find limits on that value. Remember that c is an unknown value between (in this case) 0 and 4.

$$\begin{aligned} F(4) - F(0) &= F'(c)(4-0) \quad \text{(Use the MVT on } F(4)\text{)} \\ &= \frac{1}{1+c} \cdot 4 \end{aligned}$$

We don't know what $\frac{1}{1+c}$ is, but we know that $\frac{1}{x}$ decreases from 0 to infinity, so:

$$1 = \frac{1}{1} > \frac{1}{1+c} > \frac{1}{1+4} = \frac{1}{5}.$$

Hence:

$$4 > \frac{1}{1+c} \cdot 4 > \frac{4}{5}.$$

We conclude that:

$$4 > F(4) - F(0) > \frac{4}{5}$$

and since $F(0) = 1$ we have:

$$5 > F(4) > \frac{9}{5}.$$

Our final answer is $A = \frac{9}{5}$ and $B = 5$.

Now let's compare this to what we can do with the fundamental theorem of calculus:

$$F(4) - F(0) = \int_0^4 \frac{dx}{1+x}$$

Based on what we know about the graph of $y = \frac{1}{x}$ and the area under it, we can deduce that:

$$F(4) - F(0) = \int_0^4 \frac{dx}{1+x} < \int_0^4 1dx = 4$$

and that

$$F(4) - F(0) = \int_0^4 \frac{dx}{1+x} > \int_0^4 \frac{1}{5}dx = \frac{4}{5}.$$

So once again we have:

$$\frac{4}{5} < F(4) - F(0) < 4.$$

Geometrically, we interpret $\int_0^4 \frac{dx}{1+x}$ as the area under a curve. We got an upper bound on the area by comparing it to the area of a rectangle whose height was the maximum value of $\frac{1}{1+x}$ on the interval, and got a lower bound by comparing to a rectangle whose hight was the minimum of $\frac{1}{1+x}$ on $[0, 4]$.

We could think of this as estimating $\int_0^4 \frac{dx}{1+x}$ by comparing it to two different Riemann sums, each with only *one* rectangle.

$$\text{lower Riemann sum} < \int_0^4 \frac{dx}{1+x} < \text{upper Riemann sum}$$

The Second Fundamental Theorem of Calculus

We're going to start with a continuous function f and define a complicated function $G(x) = \int_a^x f(t)\,dt$. The variable x which is the input to function G is actually one of the limits of integration. The function f is being integrated with respect to a variable t, which ranges between a and x. The variable t is a dummy variable, and is the variable of integration. Don't get t and x mixed up, even if your textbook does.

Theorem: If f is continuous and $G(x) = \int_a^x f(t)\,dt$, then $G'(x) = f(x)$.

From the point of view of differential equations, $G(x)$ solves the differential equation

$$y' = f, \qquad y(a) = 0.$$

The second fundamental theorem of calculus tells us that we can always solve this equation (by using Riemann sums if necessary).

Using The Second Fundamental Theorem of Calculus

This is the quiz question which everybody gets wrong until they practice it.

Example: Evaluate $\frac{d}{dx}\int_1^x \frac{dt}{t^2}$.

This question challenges your ability to understand what the question means. It looks very complicated, but what it really is is an exercise in recopying!

By definition, we have a function of the form $G(x) = \int_a^x f(t)\,dt$ like the one in the second fundamental theorem of calculus. The second fundamental theorem then tells us that $G'(x) = f(x)$.

So if $G(x) = \int_1^x \frac{1}{t^2}\,dt$, then $\frac{d}{dx}G(x) = \frac{1}{x^2}$.

Despite the fact that this looks like a long, elaborate problem it is actually quite easy to solve.

Question: Why don't you integrate?

Answer: Because we don't need to! When you take the derivative of something and then the antiderivative, you get back to the original thing (plus some arbitrary constant). In this case, we're taking the antiderivative and then the derivative; again we get back to the same thing. There isn't even an arbitrary constant to worry about here because the derivative of that constant is zero.

Let's see what happens when we integrate. It's slower than just recopying, but we can check it to see that the theorem does work.

$$\int_1^x t^{-2}\,dt = -t^{-1}\Big|_1^x$$

(Remember that t is our variable of integration, not x.)

$$\int_1^x t^{-2}\,dt = -\frac{1}{x} - (-1).$$

Then we take the derivative to get $\frac{d}{dx}\int_1^x \frac{dt}{t^2}$:

$$\frac{d}{dx}\left(1 - \frac{1}{x}\right) = 0 - (-x^{-2}) = \frac{1}{x^2}.$$

Proof of the Second Fundamental Theorem of Calculus

Theorem: (The Second Fundamental Theorem of Calculus) If f is continuous and $F(x) = \int_a^x f(t)\,dt$, then $F'(x) = f(x)$.

Proof: Here we use the interpretation that $F(x)$ (formerly known as $G(x)$) equals the area under the curve between a and x. Our goal is to take the derivative of F and discover that it's equal to f.

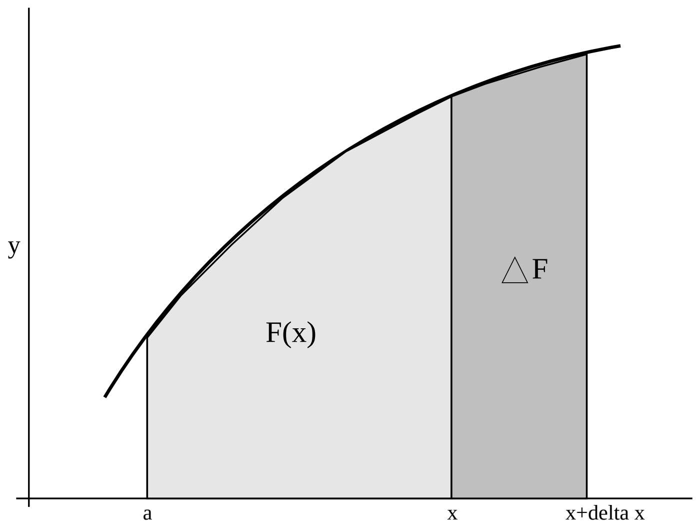


Figure 1: Graph of $f(x)$ with shaded area $F(x)$.

We graph the equation $y = f(x)$ and keep track of where a, x and $x + \Delta x$ are. This splits the area under the curve into pieces. The first piece is the area under the curve between a and x which is, by definition, $F(x)$. The second piece is a thin region; its area is ΔF, which is the change in the area under the curve as x increases by Δx.

We now approximate this thin region with area ΔF by a rectangle. Its base has width Δx and its height is close to $f(x)$ (because f is continuous). So

$$\Delta F \approx \Delta x f(x).$$

Divide both sides by Δx to get $\frac{\Delta F}{\Delta x} \approx f(x)$, then take the limit as Δx goes to zero to get the derivative:

$$F'(x) = \lim_{\Delta x \to 0} \frac{\Delta F}{\Delta x} = f(x).$$

Proof of the First Fundamental Theorem of Calculus

The first fundamental theorem says that the integral of the derivative is the function; or, more precisely, that it's the difference between two outputs of that function.

Theorem: (First Fundamental Theorem of Calculus) If f is continuous and $F' = f$, then $\int_a^b f(x)\,dx = F(b) - F(a)$.

Proof: By using Riemann sums, we will define an antiderivative G of f and then use $G(x)$ to calculate $F(b) - F(a)$.

We start with the fact that $F' = f$ and f is continuous. (It's not strictly necessary for f to be continuous, but without this assumption we can't use the second fundamental theorem in our proof.)

Next, we define $G(x) = \int_a^x f(t)\,dt$. (We know that this function exists because we can define it using Riemann sums.)

The second fundamental theorem of calculus tells us that:

$$G'(x) = f(x)$$

So $F'(x) = G'(x)$. Therefore,

$$(F - G)' = F' - G' = f - f = 0$$

Earlier, we used the mean value theorem to show that if two functions have the same derivative then they differ only by a constant, so $F - G =$ constant or

$$F(x) = G(x) + c.$$

This is an essential step in an essential proof; all of calculus is founded on the fact that if two functions have the same derivative, they differ by a constant.

Now we compute $F(b) - F(a)$ to see that it is equal to the definite integral:

$$\begin{aligned}
F(b) - F(a) &= (G(b) + c) - (G(a) + c) \\
&= G(b) - G(a) \\
&= \int_a^b f(t)\,dt - \int_a^a f(t)\,dt \\
&= \int_a^b f(t)\,dt - 0 \\
F(b) - F(a) &= \int_a^b f(x)\,dx
\end{aligned}$$

Antiderivative of $\dfrac{1}{x}$

We return to our theme of using $f = F'$ to understand F.

We recently found the antiderivative of t^{-2}. For the most part, it's easy to antidifferentiate x^n, except for the tricky case of $\frac{1}{x}$.

Example: Solve the differential equation $L'(x) = \frac{1}{x}$; $L(1) = 0$.

The second fundamental theorem of calculus tells us that the solution is:

$$L(x) = \int_1^x \frac{dt}{t} = \ln x.$$

We also know that this antiderivative must be the logarithm function, but it turns out that this way of looking at the function makes many calculations much easier.

One very interesting thing about this solution is that we started with a ratio of polynomials ($\frac{1}{x}$) and ended with a transcendental function. The natural log function can't be written in terms of our usual algebraic operations.

Area Under the Bell Curve

In addition to exotic but familiar functions like $\ln x$, we can also use definite integrals and Riemann sums to get truly *new* functions.

Example: The solution to $y' = e^{-x^2}$; $y(0) = 0$ is:

$$F(x) = \int_0^x e^{-t^2} dt$$

The graph of e^{-x^2} is known as the bell curve, and $F(x)$ describes the area under the curve. This function is extremely useful for computing probabilities.

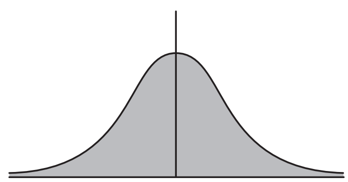


Figure 1: Graph of e^{-x^2}.

The exciting thing about $F(x)$ is that although we have a geometric definition and can compute it using Riemann sums, we can't describe it in terms of any function we've seen previously, including logarithmic and trigonometric functions. It's a completely new function. The problem of describing F is analogous to the problem of calculating the value of π — the area of a circle with radius 1. The number π is transcendental; it is not the root (zero) of an algebraic equation with rational coefficients.

Using definite integrals we can define a huge class of new functions, many of which are important tools in science and engineering.

Alternate Definition of Natural Log

The second fundamental theorem of calculus says that the derivative of an integral gives you the function back again:

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

We saw a few examples of how to use this to solve differential equations. In particular, we can solve $y' = \frac{1}{x}$ to get

$$L(x) = \int_1^x \frac{dt}{t}.$$

We could use this formula to define the logarithm function and derive all of its properties.

The first property of the function L is that $L'(x) = \frac{1}{x}$. (We defined it that way.)

The other piece of information that we need in order to completely describe the function is its output at one value; knowing its derivative only tells us its value up to a constant.

$$L(1) = \int_1^1 \frac{dt}{t} = 0$$

This will be the case with all definite integrals; if we evaluate them at their starting place, we'll get 0.

Together these two properties uniquely describe $L(x)$.

Next we want to understand the properties of the function; we'll start by graphing it. We know that the derivative of $L(x)$ is $\frac{1}{x}$; when the function is given as an integral it's easy to compute its first derivative! The derivative of L "blows up" as x approaches 0. To avoid this problem we'll consider only positive values of x.

The second derivative of L is $-\frac{1}{x^2}$. From it we learn that the graph of $L(x)$ is concave down. We know that $L(1) = 0$ and $L'(1) = 1$, which gives us a point on the graph and its slope at that point. We also know that $L'(x) > 0$ when x is positive, so we know that the function is increasing — the graph rises as we move to the right.

Knowing that $L(x)$ is increasing when x is positive allows us to work backwards from this definition to the one we used previously. If we draw the line $y = 1$ it intersects the graph of $L(x)$ at some point (we could confirm this using a Riemann sum if we had to). We'll define the number e so that this point of intersection is $(e, 1)$. In other words, e is the unique value for which $L(e) = 1$. We know there's only one such value because the graph can never "dip down" to cross the line $y = 1$ again.

Since we know $L(x)$ is increasing, we know there are no critical points; the only other interesting thing is the ends. It turns out that the limit as x

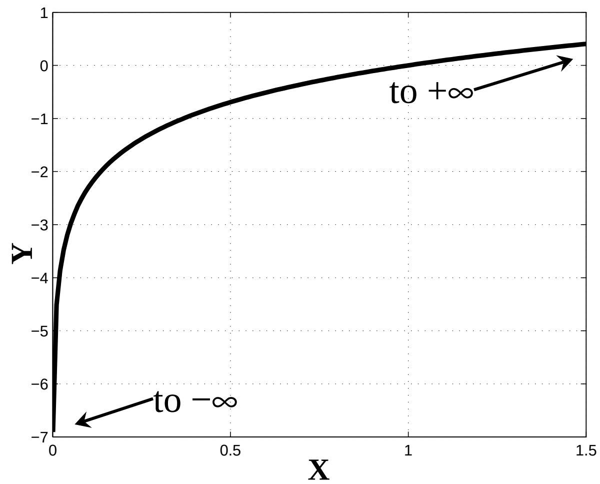


Figure 1: Graph of $y = \ln(x)$.

approaches 0 is minus infinity. As x approaches positive infinity, the limit is positive infinity (there's no horizontal asymptote). We won't get into those details today.

Instead, we'll look at one more qualitative feature of the graph; the fact that $L(x) < 0$ for $x < 1$. Why is this true?

- $L(1) = 0$ and L is increasing — if L increases to 0 as x goes toward 1 it must be negative before then.

- $L(x) = \int_1^x \frac{dt}{t} = -\int_x^1 \frac{dt}{t} < 0$ when $0 < x < 1$ because $\int_x^1 \frac{dt}{t}$ describes a positive area.

Log of a Product

Claim: $L(ab) = L(a) + L(b)$, where $L(x) = \int_1^x \frac{dt}{t}$ is an alternately defined natural log function.

To prove this, we just plug in the formula and see what happens. On the left hand side we have:

$$L(ab) = \int_1^{ab} \frac{dt}{t} = \int_1^a \frac{dt}{t} + \int_a^{ab} \frac{dt}{t}$$

By definition, $\int_1^a \frac{dt}{t} = L(a)$. If we could show that $\int_a^{ab} \frac{dt}{t} = L(b)$, we'd be done with the proof.

It turns out that we can prove this by using a change of variables. We start with $\int_a^{ab} \frac{dt}{t} = L(b)$, and substitute $t = au$ (so $dt = a\,du$). The limits of integration are from $u = 1$ to $u = b$. If we plug these into $\int_a^{ab} \frac{dt}{t}$, we get:

$$\int_a^{ab} \frac{dt}{t} = \int_{u=1}^{u=b} \frac{a\,du}{au} = \int_1^b \frac{du}{u} = L(b).$$

We can now conclude that:

$$L(ab) = L(a) + L(b)$$

The Area Under the Bell Curve

Our last example of a definite integral is:

$$F(x) = \int_0^x e^{-t^2}\, dt.$$

As we've already remarked, this is a new function that we can't express in terms of functions we already know. To begin to understand the properties of this function F we'll sketch its graph.

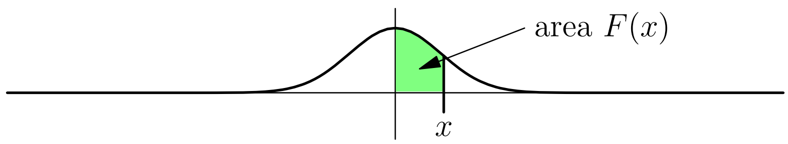


Figure 1: Area under e^{-t^2}.

The fundamental theorem tells us right away that:

$$F'(x) = e^{-x^2}.$$

It's easy to find the value of $F(0)$, because the "starting place" of the integral is 0.

$$F(0) = \int_0^0 e^{-t^2}\, dt = 0.$$

We can also compute the second derivative:

$$F''(x) = -2xe^{-x^2}$$

The first derivative is always positive, so $F(x)$ is always increasing. Because the sign of $-2xe^{-x^2}$ is just the sign of $-2x$, its graph will be concave down when $x > 0$ and concave up when $x < 0$. And finally, $F'(0) = e^{-0^2} = 1$. (We often define functions so that their graphs have slope 1 in convenient locations.)

By combining all this information we get a pretty good idea of what the graph of the function looks like, even though we cannot write down a strictly algebraic equation for $F(x)$.

We want to know as much as possible about this function, so we'll discuss a few more features before moving on.

First we show that $F(x)$ is odd. As mentioned in Figure 1, $F(x)$ is the area under the graph of e^{-t^2} between 0 and x. From the symmetry of the graph we see:

$$\int_{-x}^0 e^{-x^2}\, dx = \int_0^x e^{-x^2}\, dx,$$

so:

$$F(-x) = \int_0^{-x} e^{-x^2}\, dx = -\int_{-x}^0 e^{-x^2}\, dx = -\int_0^x e^{-x^2}\, dx = -F(x).$$

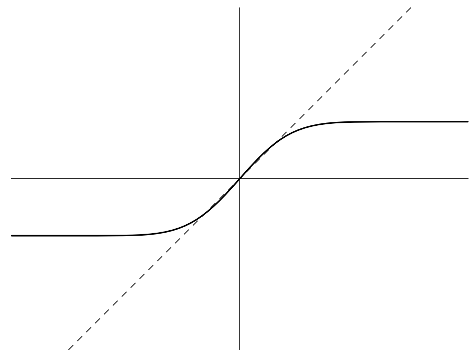


Figure 2: Graph of $F(x) = \int_0^x e^{-t^2}\,dt$.

We conclude that $F(x)$ is an odd function.

Because F is odd we know that the part of its graph to the right of the y-axis is exactly a rotation of the part of its graph to the left of the y-axis. If we know the shape of one branch we immediately know the shape of the other branch.

Our final step to understanding the graph is to figure out what happens at the ends. It turns out that the graph approaches a horizontal asymptote as x approaches positive infinite and, by symmetry, as x goes to negative infinity. On the right, the graph rises to a certain level; on the left it falls to the negative of that level.

What is that level? It's the area under the graph of $f(x) = e^{-t^2}$ between 0 and infinity: $\frac{\sqrt{\pi}}{2}$. It took several years for mathematicians to discover the exact value of this area.

$$\lim_{x\to\infty} F(x) = \frac{\sqrt{\pi}}{2}$$

$$\lim_{x\to-\infty} F(x) = -\frac{\sqrt{\pi}}{2}$$

Because 1 is a nicer number to work with than $\frac{\sqrt{\pi}}{2}$, people define the *error function* to be:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt = \frac{2}{\sqrt{\pi}} F(x).$$

This is a famous function, as is its cousin the standard normal distribution.

More New Functions from Old

Using definite integrals, we can define many transcendental or "new" functions which cannot be expressed in elementary terms.

Example: Fresnel Integrals

$$C(x) = \int_0^x \cos(t^2)\, dt$$
$$S(x) = \int_0^x \sin(t^2)\, dt$$

These are named after Augustin-Jean Fresnel and are used in optics.

Example: From Fourier Analysis

$$H(x) = \int_0^x \frac{\sin t}{t}\, dt$$

Example: Logarithmic Integral

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\ln t}$$

This function is significant because it is approximately equal to the number of primes less than x. If you can describe precisely how close the value of $\mathrm{Li}(x)$ is to the exact number of primes less than x you'll have proven the Riemann hypothesis; a task mathematicians have been working on for over a century.

Question: Are we supposed to understand this stuff?

Answer: That's a good question. We're going to see a lot more of the function $F(x) = \int_0^x e^{\ \ t}\ dt$ because it's associated with the normal distribution. The functions described in this segment are simply examples of other transcendental functions that are important and described by definite integrals. For this class, the only thing that you'll need to do with such functions are things like understanding the derivative, the second derivative, and sketching their graphs.

Areas Between Curves

Suppose you have two curves, $y = f(x)$ above and $y = g(x)$ below. You want to find the area between the two curves bounded on the left by $x = a$ and on the right by $x = b$.

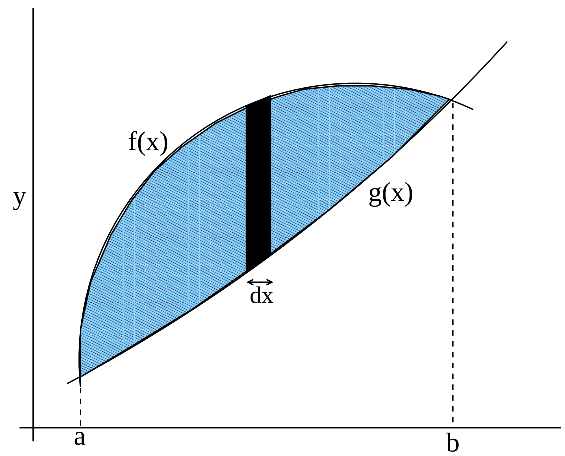


Figure 1: Finding the area between two intersecting functions.

As we did with Riemann sums, we can (approximately) chop this area up into thin rectangles. Each rectangle will have width dx and height $f(x) - g(x)$, so will have area

$$\underbrace{(f(x) - g(x))}_{\text{height}} \underbrace{dx}_{\text{base}}.$$

In order to get the whole area, sum the areas of all these rectangles:

$$A = \int_a^b (f(x) - g(x))\, dx.$$

There are two key steps to solving problems with integrals. The first is figuring out what to integrate. The function being integrated is called the *integrand*. The second is finding the *limits of integration* — in this case a and b. Once we have these we can compute the integral, either numerically or by finding an antiderivative. Without the integrand and limits of integration we can't find the value of the integral.

Example: Find the area between $x = y^2$ and $y = x - 2$

First, graph these functions. If skip this step you'll have a hard time figuring out what the boundaries of your area is, which makes it very difficult to compute the area!

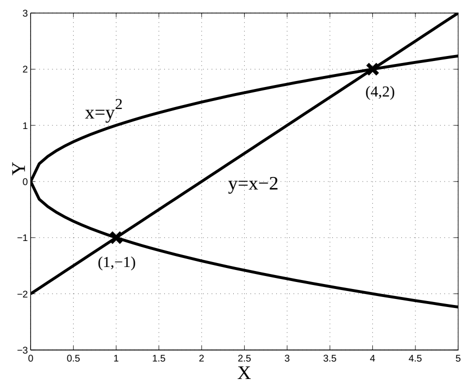


Figure 1: Finding the area between two intersecting graphs.

There are two ways of finding the area between these two curves: the hard way and the easy way.

Hard Way: slice it vertically

First, we'll try chopping the region up into vertical rectangles.

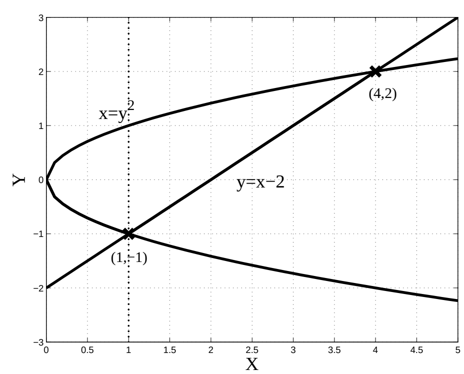


Figure 2: The area between $x = y^2$ and $y = x - 2$ split into two subregions.

If we slice the region between the two curves this way, we need to consider two different regions. Where $x > 1$, the region's lower bound is the straight line. For $x < 1$, however, the region's lower bound is the lower half of the sideways parabola. The left end of the region is at $x = 0$.

We must find the crossing points of the two curves; in other words, we find the values of x and y that satisfy both equations simultaneously.

$$x = y^2 \quad \text{and} \quad y = x - 2$$

so:

$$\begin{aligned} y - 2 &= y^2 \\ y^2 - y - 2 &= 0 \\ (y-2)(y+1) &= 0 \end{aligned}$$

We conclude that:

$$y = 2 \quad \text{or} \quad y = -1.$$

We can plug these values of y back in to either equation to find the associated x values:

$$\begin{aligned} y &= x - 2 \\ 2 &= x - 2 \\ 4 &= x. \end{aligned}$$

If we perform a similar equation with $y = -1$ we'll find that the two points of intersection are $(1, -1)$ and $(4, 2)$.

The equation of the upper half of the sideways parabola is $y = \sqrt{x}$ and that of the lower half is $y = -\sqrt{x}$. The equation of the lower right hand boundary of the region is just $y = x - 2$.

We find the area A between the two curves by integrating the difference between the top curve and the bottom curve in each region:

$$A = \underbrace{\int_0^1 (\overbrace{\sqrt{x}}^{\text{top}} - \overbrace{(-\sqrt{x})}^{\text{bottom-l}}\,))\,dx}_{\text{left}} + \underbrace{\int_1^4 (\overbrace{\sqrt{x}}^{\text{top}} - \overbrace{(x-2)}^{\text{bottom-r}}\,)\,dx}_{\text{right}}$$

The rest of this calculation is easy; just evaluate the integrals.

$$\begin{aligned} A &= 2\int_0^1 \sqrt{x}\,dx + \int_1^4 (-x + \sqrt{x} + 2)\,dx \\ &= 2\left[\frac{2}{3}x^{3/2}\right]_0^1 + \left[-\frac{1}{2}x^2 + \frac{2}{3}x^{3/2} + 2x\right]_1^4 \\ &= 2\left(\frac{2}{3} - 0\right) + \left(-\frac{4^2}{2} + \frac{2}{3}\cdot 4^{3/2} + 8\right) - \left(-\frac{1}{2} + \frac{2}{3} + 2\right) \\ &= \frac{4}{3} - 8 + \frac{16}{3} + 8 + \frac{1}{2} - \frac{2}{3} - 2 \\ A &= \frac{9}{2}. \end{aligned}$$

Easy Way: Slice it horizontally

There's a much quicker way to complete this area calculation; you should look for an easier way as soon as you notice the need to split the region into parts. The quicker way is similar in principle but reverses the roles of x and y; in this method we slice the area in question into horizontal rectangles.

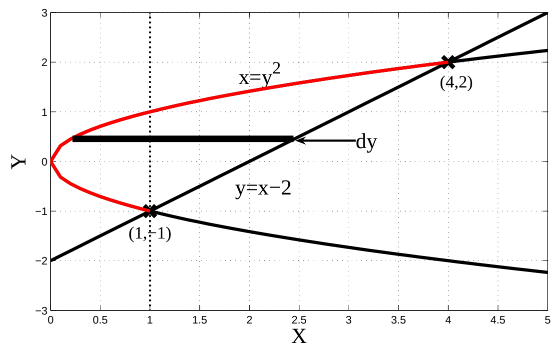


Figure 3: The area between $x = y^2$ and $y = x - 2$ and one horizontal rectangle.

The height of these rectangles is dy; we get their width by subtracting the x-coordinate of the edge on the left curve from the x-coordinate of the edge on the right curve. (If you get mixed up and subtract the right from the left you'll get a negative answer.) The left curve is the sideways parabola $x = y^2$. The right curve is the straight line $y = x - 2$ or $x = y + 2$.

The limits of integration come from the points of intersection we've already calculated. In this case we'll be adding the areas of rectangles going from the bottom to the top (rather than left to right), so from $y = -1$ to $y = 2$.

$$\begin{aligned} A &= \int_{y=-1}^{y=2} \left[(y+2) - y^2\right] dy \\ &= \left[\frac{-y^3}{3} + 2y + \frac{y^2}{2}\right]_{-1}^{2} \\ &= \left(\frac{4}{2} + 4 - \frac{8}{3}\right) - \left(\frac{1}{2} - 2 + \frac{1}{3}\right) \\ A &= \frac{9}{2} \end{aligned}$$

You'll notice that if you plug the limits of integration into the integrand, you get 0. This makes sense; as y goes toward -1 and 2 the width of the rectangles approaches 0.

Volumes by Slicing

Suppose you have a loaf of bread and you want to find the volume of the loaf. One way to do this is to find the volume of each slice and then add up their volumes.

The volume of a slice of bread is its thickness dx times the area a of the face of the slice (the part you spread butter on). So $\Delta V \approx A\Delta x$. In the limit, $dV = A(x)\,dx$. (If your loaf of bread is not perfectly regular, the area of a face might change from slice to slice.) To get the entire volume, sum the volumes of all the slices:

$$V = \int A(x)\,dx$$

The Riemann sum approximating this volume looks like $\displaystyle\sum_{i=1}^{n} A_i\,\Delta x$ if the loaf has n slices.

Solids of Revolution

In theory we could take any three dimensional object and estimate its volume by slicing it into slabs and adding the volumes of the slabs. In practice we'll concentrate exclusively on *solids of revolution*. These are formed by taking an area — for example the arc over the x-axis shown in Figure 1 — and revolving it about an axis to see what volume it sweeps out. If you rotate that arc and its interior about the x-axis you get a shape like an American football or a rugby ball. (See Figure 2.)

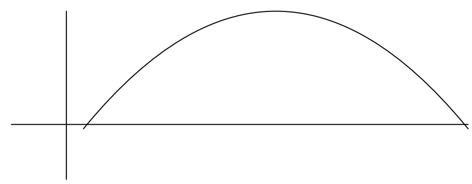


Figure 1: An arc over the x-axis.

Method of Disks

How would we slice up this ball to find its volume? We'll start with the two dimensional picture of an arc over the x-axis. Two dimensional figures are much easier to draw and understand than three dimensional ones; when possible you should avoid drawing three dimensional figures. In this picture we draw a thin rectangle whose base lies on the x-axis and whose height is the height of the arc. The width of this rectangle is dx.

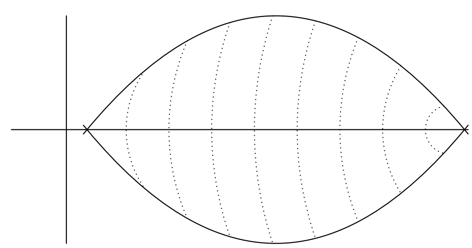


Figure 2: The solid formed by revolving an arc about the x-axis.

Next we need to visualize how this rectangle is related to the three dimensional volume of revolution. When we rotate the arc about the x-axis, the rectangle rotates also as if it were hinged. As it rotates it sweeps out a disk or coin shape. This corresponds to one "slice" of our solid ball, like a slice of bread. The method that we're describing for figuring out the volume of the ball is called the *method of disks* because we're slicing the ball into disks.

We add the volumes of these disks to find the volume of the ball. If the height of the arc over the x-axis is y, then the area $A(x)$ of a face of the disk or slice is πy^2 because the rectangle swept out a circular shape as it spun about the x-axis. The thickness of the disk is dx, so the volume of the disk is

$$dV = (\pi y^2)\, dx.$$

This will be the integrand in the formula for volume associated with the method of disks.

$$V = \int (\pi y^2)\, dx$$

Notice that we have both a y and a dx in our formula, yet we don't have a formula describing y in terms of x. That formula depends on the equation $y = f(x)$ of the arc over the x-axis, and will change depending on the situation.

In addition, we haven't specified any limits of integration yet; again those will depend on the situation.

Example: Volume of a Sphere

Let's practice the method of disks by finding the volume of a soccer ball. We'll find the volume of revolution formed by rotating a circle with center at $(a, 0)$ and radius a about the x-axis. This will sweep out a ball of radius a. Our goal is to find out the volume of this ball.

Remember that our integrand will be $dV = \pi y^2\, dx$. In order to use this we need a formula for y in terms of x, so we need the equation of the circle of radius a centered at $(a, 0)$:

$$(x-a)^2 + y^2 = a^2.$$

We need to describe y in terms of x:

$$\begin{aligned}
(x-a)^2 + y^2 &= a^2 \\
y^2 &= a^2 - (x-a)^2 \\
&= a^2 - (x^2 - 2ax + a^2) \\
y^2 &= 2ax - x^2
\end{aligned}$$

We can stop here (without taking any square roots) because the value we need to plug into our formula is y^2. We get:

$$V = \int \pi(2ax - x^2)\, dx.$$

Of course we can't evaluate this integral without knowing the limits of integration. Luckily those limits are easy to find in this example. We're integrating with respect to x. The lowest value of x in our circle is $x = 0$ and the highest is $2a$, so x ranges between 0 and $2a$. Our formula for the volume of a ball becomes:

$$\begin{aligned}
V &= \int_0^{2a} \pi(2ax - x^2)\, dx \\
&= \pi\left(ax^2 - \frac{x^3}{3}\right)\Bigg|_0^{2a} \\
&= \pi\left(4a^3 - \frac{8a^3}{3}\right) - 0 \\
&= \left(\frac{12}{3} - \frac{8}{3}\right)\pi a^3 \\
V &= \frac{4}{3}\pi a^3
\end{aligned}$$

One nice thing about this formula is that we've found more than just the volume of the ball. If we change the upper limit of integration we can also find the volume of a piece sliced off the ball.

$$V(x) \quad = \quad \text{Volume of a chopped off portion of the sphere with width } x$$

$$\begin{aligned} &= \int_0^x \pi(2at - t^2)\, dt \\ V(x) &= \pi\left(ax^2 - \frac{x^3}{3}\right) \end{aligned}$$

If we plug in $x = a$ we should get half the volume of the sphere:

$$\begin{aligned} V(a) &= \text{Volume of a half sphere} \\ &= \pi\left(a^3 - \frac{a^2}{3}\right) \\ &= \pi\left(a^3 - \frac{a^2}{3}\right) \\ &= \pi\frac{2}{3}a^3 \\ &= \frac{1}{2}\left(\frac{4}{3}\pi a^3\right) \end{aligned}$$

This is a good way to check our work.

The formula $V(x) = \pi\left(ax^2 - \frac{x^3}{3}\right)$ turns out to be useful in predicting the behavior of particles in a fluid. When large spherical particles are being pushed around by small ones, will they tend to cluster together or to stick to the sides of the container? Finding the answer to this question involves adding the volumes of two slices off a sphere to find the volume of a lens shaped region.

Example: Volume of a Cauldron

In our next, Halloween themed, example we'll compute the volume of the region shown below.

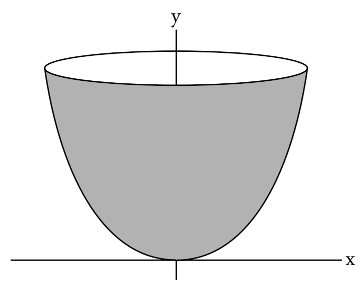


Figure 1: $y = x^2$ rotated around the y-axis.

We could use the method of disks to calculate this volume, but instead we will use the other standard method of finding volumes — the *method of shells*.

A cross section of the cauldron has boundaries $y = x^2$ and $y = a$. We revolve this cross section around the y-axis to get our cauldron. (Note that we can revolve shapes around the y-axis as well as the x-axis.) Our "shell" will be the result of revolving a thin rectangle with its base on $y = x^2$ and its top at $y = a$, as shown in Figure 2. This "shell" shape might also be described as a cylinder. Make your own by rolling a piece of piece of paper into a tube!

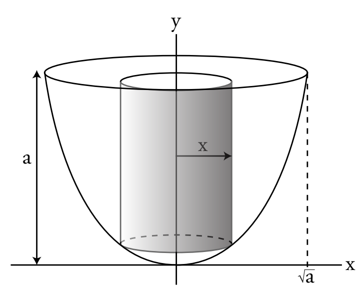


Figure 2: Cylindrical shell with radius x, thickness dx and height $a - y = a - x^2$.

We need to compute the volume of this shell. It's thickness is dx, and its height is $y_{top} - y_{bottom} = a - x^2$. Since the shell is very thin, we get a good approximation of its volume by "unrolling" it like a piece of paper and computing the volume of a rectangular slab with thickness dx, height $a - x^2$ and width equal to the circumference of the shell. To compute the circumference we multiply the radius r by 2π; our estimate of the shell's volume is then:

$$dV = (2\pi x)(a - x^2)\,dx = 2\pi(ax - x^3)\,dx$$

(We can roughly check our work by noting that we're multiplying three lengths here, so the units do match up on both sides of the equation.)

To compute the volume of the cauldron we'll integrate this; all that's left is to find the limits. The entire volume of the cauldron is swept out by the right side of the parabola as it spins about the y-axis, so our limits of integration start at 0 (not $-\sqrt{a}$). In other words, if we just rotate the right half of this it covers the left half; if we counted the volume swept out by the left half we would be counting the volume twice.

The upper limit of integration is at the farthest rightmost spot, where $y = a$ and $y = x^2$ simultaneously; in other words, at $x = \sqrt{a}$. Getting the correct limits is as important as getting the right integrand.

So the volume of the cauldron is:

$$\begin{aligned} V &= 2\pi \int_0^{\sqrt{a}} (ax - x^3)\, dx \\ &= 2\pi \left(a\frac{x^2}{2} - \frac{x^4}{4} \right) \Big|_0^{\sqrt{a}} \\ &= 2\pi \left(\frac{a^2}{2} - \frac{a^2}{4} \right) = 2\pi \left(\frac{a^2}{4} \right) \\ V &= \frac{\pi}{2} a^2. \end{aligned}$$

Question: How do you know whether the rectangle should be vertical or horizontal?

Answer: You can always set it up both ways. One way may be a difficult calculation and one way may be an easier calculation. In the example of $y = x^2$ and $y = x - 2$ the horizontal and the vertical calculations were quite different in character. One of them was really a mess, and one of them was a little easier. This is often the case, and once in a while, one of them is impossible and the other one is possible. By choosing your method carefully you can save yourself a lot of work.

Warning about units.

Previously, we calculated the volume of a parabolic "cauldron" to be $\frac{\pi}{2}a^2$. There's something fishy about this expression — it looks as if it has units of area, but it's describing a volume. In general, we must be very aware of what units we're using.

Suppose the height of the cauldron is $a = 100$cm. Then:

$$\begin{aligned} V &= \frac{\pi}{2}(100)^2\,\text{cm}^3 \\ &= \frac{\pi}{2}10^4\,\text{cm}^3 \\ &= \frac{\pi}{2}10 \sim 16\,\text{liters} \end{aligned}$$

Next, suppose that the height of the cauldron is $a = 1$m. Then:

$$\begin{aligned} V &= \frac{\pi}{2}(1)^2\,\text{m}^3 \\ &= \frac{\pi}{2}10^6\,\text{cm}^3 \\ &= \frac{\pi}{2}1000 \sim 1600\,\text{liters} \end{aligned}$$

But 100cm = 1m. Why are the answers different?

The problem is that we don't know the units in the equation $y = x^2$. If the units are centimeters, then 100cm $= 10^2$cm. If the units are meters then 1m $= 1^2$m. When we use centimeters as units, the cauldron is five times as tall as it is wide, so it looks like:

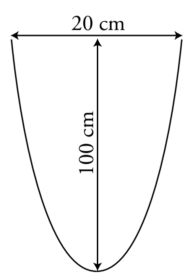


Figure 1: Cauldron cross section for units of centimeters.

When we interpret $y = x^2$ in meters, we find that the cauldron is twice as wide as it is tall, which seems more likely in the context of the problem.

This confusion about units arose because the equation $y = x^2$ is not scale-invariant.

Average Value

You already know how to take the average of a finite set of numbers:

$$\frac{a_1 + a_2}{2} \text{ or } \frac{a_1 + a_2 + a_3}{3}$$

If we want to find the average value of a function $y = f(x)$ on an interval, we can average several values of that function:

$$\text{Average} \approx \frac{y_1 + y_2 + ... + y_n}{n}.$$

As was mentioned previously, if we let the number of values n approach infinity we get:

$$\text{Continuous Average} = \frac{1}{b-a}\int_a^b f(x)\,dx = \text{Ave}(f).$$

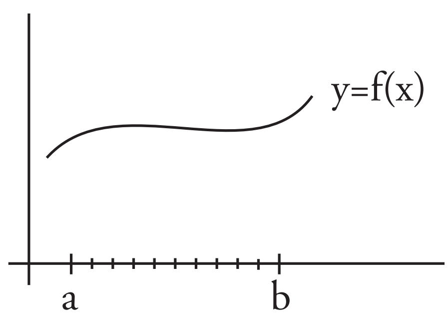


Figure 1: $a \le x \le b$.

Why does this describe the average value of $f(x)$? Imagine that you have $n+1$ equally spaced points $a = x_0 < x_1 < x_2 < ... < x_n = b$. The distance between each pair of points is $\Delta x = \frac{b-a}{n}$. Let $y_0 = f(x_0)$, $y_1 = f(x_1), ...y_n = f(x_n)$.

Then the Riemann sum approximating the area under the curve is:

$$(y_1 + y_2 + \cdots + y_n)\Delta x.$$

As n approaches infinity this approaches the area under the curve, which is:

$$\int_a^b f(x)\,dx.$$

$$\begin{aligned}
\frac{1}{b-a}\int_a^b f(x)\,dx &\approx \frac{1}{b-a}(y_1+y_2+\cdots+y_n)\Delta x \\
&= \frac{1}{b-a}(y_1+y_2+\cdots+y_n)\frac{b-a}{n} \\
&= \frac{y_1+y_2+\cdots+y_n}{n},
\end{aligned}$$

so:

$$\frac{1}{b-a}\int_a^b f(x)\,dx \approx \frac{y_1+y_2+\cdots+y_n}{n}.$$

The only difference between the average value and the integral (area under the curve) is that we're dividing by the length of the interval.

Example: Find the average value of $f(x) = c$ on the interval $[a, b]$, where a, b and c are arbitrary constants.

$$\begin{aligned}
\frac{1}{b-a}\int_a^b c\,dx &= \frac{1}{b-a}\cdot(\text{Area of a } (b-a) \text{ by } c \text{ rectangle}) \\
&= \frac{1}{b-a}\cdot(b-a)\cdot c \\
&= c
\end{aligned}$$

If the value of $f(x)$ is always c, then the average value of $f(x)$ had better be c. This confirms that our formula for the average value of a function works, and in particular it confirms that $\frac{1}{b-a}$ is the correct normalizing factor. In this case our Riemann sum becomes:

$$\begin{aligned}
\frac{y_1+y_2+\cdots+y_n}{n} &= \frac{\overbrace{c+c+\cdots+c}^{n\text{times}}}{n} \\
&= \frac{nc}{n} \\
&= c
\end{aligned}$$

and we see why we needed the n in the denominator.

Average Height

Find the average height of a point on a unit semicircle.

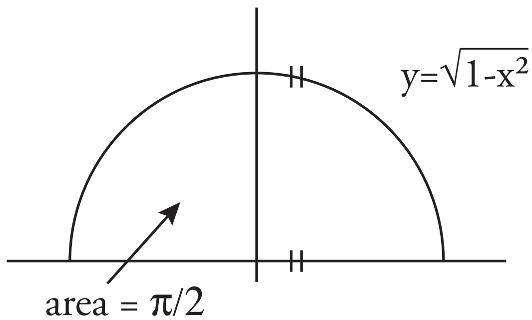


Figure 1: The unit semicircle and an interval dx.

Here $f(x) = \sqrt{1-x^2}$ for $-1 \le x \le 1$, so $a = -1$ and $b = 1$. The average value of $f(x)$ is:

$$\begin{aligned}
\text{Avg}(f) &= \frac{1}{b-a}\int_a^b f(x)\,dx \\
&= \frac{1}{2}\int_{-1}^{1} \sqrt{1-x^2}\,dx \\
&= \frac{1}{2}(\text{Area of a unit semicircle}) \\
&= \frac{1}{2}\left(\frac{\pi}{2}\right) \\
&= \frac{\pi}{4}.
\end{aligned}$$

(We will eventually learn how to find the antiderivative of $\sqrt{1-x^2}$ in the unit on techniques of integration.)

Average with Respect to Arc length

Find the average height of a point on a unit semicircle *with respect to the arc length* θ.

When taking averages, it's extremely important to specify the variable with respect to which the average is taking place. The answer may be different depending on the variable!

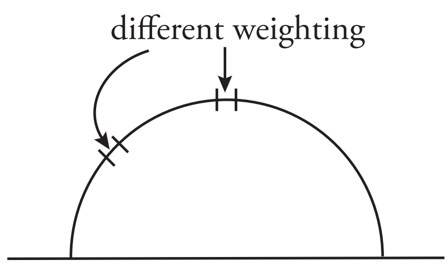


Figure 1: Equal arc lengths correspond to different distances on the x-axis.

As you can see from Figure 1, equal distances along the arc of the semicircle overshadow different lengths on the x-axis. Taking the average with respect to θ will weight the lower parts of the semicircle more heavily than the higher ones. We expect the average with respect to arc length be less than $\frac{\pi}{4}$.

Then average is still given by $\frac{1}{b-a}\int_a^b f(\theta)d\theta$. This time, $a = 0$ and $b = \pi$. The integrand is $y = \sin\theta$, which is the height of the semicircle in terms of θ. So our average height is:

$$\begin{aligned} \frac{1}{b-a}\int_a^b f(\theta)\,d\theta &= \frac{1}{\pi}\int_0^\pi \sin\theta\,d\theta \\ &= \frac{1}{\pi}(-\cos\theta)\Big|_0^\pi \\ &= \frac{1}{\pi}\left(-\cos\pi - (-\cos 0)\right) \\ &= \frac{2}{\pi} \end{aligned}$$

Let's see if we can check our work. The average height with respect to arc length is $\frac{2}{\pi}$. The average height with respect to horizontal distance is $\frac{\pi}{4}$. Points on the semicircle with lower height "count for more" in the computation with respect to arc length, so we'd expect a lower average. This does turn out to be the case because:

$$\frac{2}{\pi} < \frac{\pi}{4} \quad \text{if and only if} \quad 8 < \pi^2.$$

Since $\pi^2 > 9$ we have reason to believe we've found the right answer

Question: How do we interpret the result of the average with respect to arc length?

Answer: One way of thinking of it anticipates our next subject, which is probability. Suppose you picked a point at random along the base of the semicircle (with equal likelihood between -1 and 1) and checked the height above that point. The expected value of that height is given by the first calculation: $\frac{\pi}{4}$.

The second calculation tells you the expected value of the height of a point picked at random on the semicircle, if you were equally likely to pick any point on the semicircle.

Those two average heights are different because distance along the semicircle is different from distance along the x-axis.

Question: Shouldn't the average with respect to arc length have a *bigger* value because the arc length is *longer*?

Answer: Averages never work that way; when we multiply by $\frac{1}{b-a}$ we are dividing by the total length.

However, the average of a constant is that same constant regardless of what variable we use because we compensate by dividing by $b - a$. The difference between an integral and an average is that we're dividing by that total.

Weighted Averages

A *weighted average* is calculated by dividing the weighted total value of a fraction by the total of the weighting function:

$$\frac{\int_a^b f(x)w(x)\,dx}{\int_a^b w(x)\,dx}.$$

Multiplying by $w(x)$ makes some values of $f(x)$ contribute more to the total than other values, depending on the value of x and $w(x)$. Dividing by the integral of $w(x)$ is analogous to dividing by the length or by the number of values.

First we check that this makes sense by confirming that the weighted average of a constant is that same constant:

$$\frac{\int_a^b cw(x)\,dx}{\int_a^b w(x)\,dx} = \frac{c\int_a^b w(x)\,dx}{\int_a^b w(x)\,dx} = c.$$

We see that we were correct to put $\int_a^b w(x)\,dx$ in the denominator.

Now pretend you have a stock which you bought for \$10 one year. Six months later you brought some more for \$20, and then you bought some more for \$30. What's the average price of your stock?

It depends on how many shares you bought. If you bought w_1 shares the first time, w_2 shares the second time and w_3 shares the third time, the total amount that you spent is

$$10w_1 + 20w_2 + 30w_3.$$

The average price per share is the total price divided by the total number of shares:

$$\frac{10w_1 + 20w_2 + 30w_3}{w_1 + w_2 + w_3}$$

This is the discrete analog of the continuous average

$$\frac{\int_a^b f(x)w(x)\,dx}{\int_a^b w(x)\,dx}.$$

The function f is the function describing the price of a share and the weights are the amounts (relative importance) of the different purchases.

Question: You can't factor out the $f(x)$, can you?

Answer: When we found the weighted average of a constant, we factored out c. In

$$\frac{\int_a^b f(x)w(x)\,dx}{\int_a^b w(x)\,dx}$$

we cannot factor out $f(x)$. If the weighted average is interesting you have to do two different integrals to calculate it. It's only when $f(x)$ is constant that you can factor it out (in which case, the calculation is not very interesting at all).

Boiling Cauldron: Introduction

Now let's fill the cauldron from our example with water and light a fire under it to get the water to boil (at 100°C). Let's say it's a cold day: the temperature of the air outside the cauldron is 0^oC. How much energy does it take to boil this water, i.e. to raise the water's temperature from 0°C to 100°C?

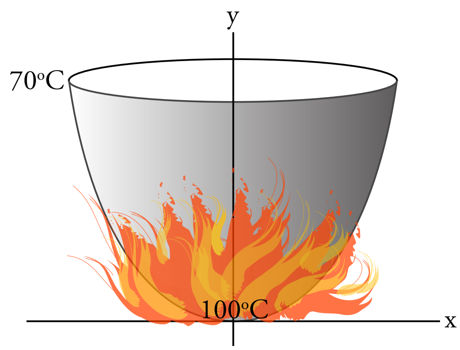


Figure 1: The boiling cauldron.

The temperature of the water is not the same at each level in the kettle. At the bottom of the kettle, where you're heating it up, it's at its highest temperature — 100 degrees Celsius. At the top, it's going to be, say, 70 degrees Celsius. The temperature is varying in height; if it varies linearly the temperature at height y will be $100 - \frac{30}{a}y$ degrees Celsius.

The total amount of heat you need to add is going to be temperature times volume, and some places will get more heat than others. The base of the cauldron will be hottest, but it also has the least volume. The cauldron is widest at the top, so we have more water to heat at that level.

At each horizontal level, the temperature is constant, so we'll use horizontal rectangles in this calculation. We'll revolve these short horizontal rectangles about the y-axis to get disks, calculate the amount of heat needed for each disk, then integrate that value with respect to dy.

Boiling Cauldron, Continued

Last class we asked how much energy it would take to boil water in a cauldron whose shape is found by rotating a parabola with height 1 meter and width 2 meters about the y-axis. The approximately 1600 liters of water would start at a uniform $T = 0$ degrees Celsius. We assumed that the final temperature would be given by the formula

$$T = 100 - 30y\text{degrees Celsius.}$$

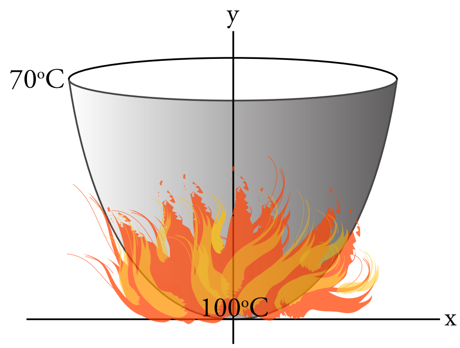


Figure 1: The boiling cauldron.

It's important to realize T is constant on each horizontal level and so the method of disks is the best way to add up the energy needed. If we set up the integral with shells, T would vary from top to bottom of the shell and we'd have to solve a second calculus problem to figure out how much heat was needed for that shell.

The equation of the parabola is $y = x^2$, so each disk will have radius $x = \sqrt{y}$ and height dy. To calculate the amount of energy needed we'll have to add up

$$\text{energy} = \text{degrees} \cdot \text{volume}.$$

So our answer will be:

$$\int_0^1 \underbrace{(100 - 30y)}_{\text{temperature}} \underbrace{\pi x^2 \, dy.}_{\text{disk volume}}$$

$$\begin{aligned}
\int_0^1 (100 - 30y)\pi y \, dy &= \int_0^1 (100y - 30y^2)\pi \, dy \\
&= \left[50\pi y^2 - 10\pi y^3\right]_0^1 \\
&= 50\pi - 10\pi \\
&= 40\pi \deg \cdot \mathrm{m}^3
\end{aligned}$$

Our answer has units of degrees Celsius times cubic meters. Let's translate that into something more familiar.

$$\begin{aligned} 40\pi \deg \cdot \mathrm{m}^3 \left(\frac{1\mathrm{cal}}{\deg \cdot \mathrm{cm}^3}\right)\left(\frac{100\mathrm{cm}}{\mathrm{m}}\right)^3 &= 40\pi \cdot 10^6 \text{ calories} \\ &= 40\pi \cdot 1000 \,\mathrm{Kcal} \\ &\approx 125,000\pi \,\mathrm{Kcal} \end{aligned}$$

One candy bar has about 250 Kcal, so it takes the energy of about 500 candy bars to heat the cauldron.

Question: What does this integral give us?

Answer: We must heat each milliliter of water in the cauldron up to some temperature between 70 and 100 degrees Celsius. For each milliliter there's some amount of energy needed to do that — water lower down in the pot needs to be heated to a higher temperature, and so will require more energy. A calorie is the amount of energy needed to raise one cubic centimeter of water by one degree Celsius. This integral adds up the energy needed to heat every single drop of water in the cauldron to exactly the right temperature.

Before we go on, let's compute the average final temperature. (The average initial temperature is 0 because the temperature is initially constant.) The formula for a weighted average is

$$\frac{\int_a^b f(x)w(x)\,dx}{\int_a^b w(x)\,dx}.$$

In this example, our function $f(x)$ is the temperature T and our weight $w(y) = \pi y$ corresponds to the volume of water in a horizontal disk; the denominator is the total volume of water in the cauldron. The limits of integration are still 0 and 1.

$$\begin{aligned} \frac{\int_0^1 T\pi y\,dy}{\int_0^1 \pi y\,dy} &= \frac{40\pi}{\pi/2} \\ &= 80 \text{ degrees} \end{aligned}$$

The value of the weight function $w(y)$ is different at different heights. This makes sense; there's more water at the top of the cauldron than at the bottom. If you tried to take the average the ordinary way you would get:

$$\frac{T_{\max} + T_{\min}}{2} = \frac{100 + 70}{2} = 85 \text{ degrees}.$$

This estimate is higher than the weighted average because it doesn't take into account that there is more water at the top of the cauldron (70 degrees) than at the bottom (100 degrees).

Probability Example

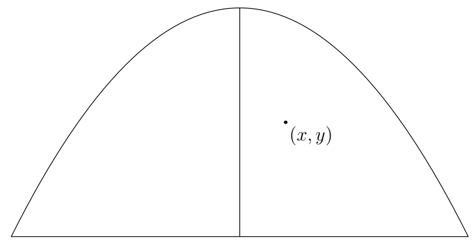


Figure 1: Choose a point at random.

Probability, volumes and weighted averages are three of the most important applications of integration. We'll analyze the probability experiment of picking a point "at random" in the region bounded below by $y = 0$ and above by $y = 1 - x^2$. Inside this parabolic region, the probability of picking a point in a given location is proportional to the area of the location.

What is the chance that $x > 1/2$? In other words, for a point picked at random, what is the probability that $x > 1/2$? Or, what is $P(x > 1/2)$?

$$\begin{aligned} \text{Probability} &= \frac{\text{Part}}{\text{Whole}} \\ &= \frac{\text{Target Area}}{\text{Entire Area}} \\ &= \frac{\text{Success}}{\text{All Possibilities}} \end{aligned}$$

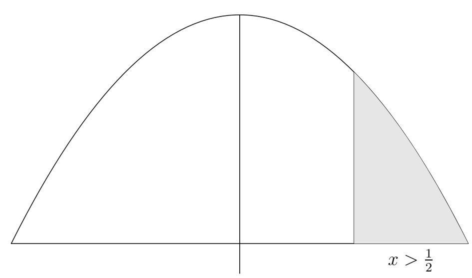


Figure 2: What is the probability that $x > \frac{1}{2}$?

The probability will just be the ratio of the two areas:

$$\frac{\int_{1/2}^{1} (1 - x^2)\, dx}{\int_{-1}^{1} (1 - x^2)\, dx}.$$

If we like, we can think of this as a weighted average with $w(x) = 1-x^2$, $a = -1$, $b = 1$ and:

$$f(x) = \begin{cases} 0 & \text{when } x < 1/2 \\ 1 & \text{when } x \geq 1/2. \end{cases}$$

$$\begin{aligned} P(x > 1/2) &= \frac{\int_{1/2}^{1} (1-x^2)\,dx}{\int_{-1}^{1} (1-x^2)\,dx} \\ &= \frac{(x - \frac{x^3}{3})\Big|_{1/2}^{1}}{(x - \frac{x^3}{3})\Big|_{-1}^{1}} \\ &= \frac{\left(\frac{2}{3} - \frac{11}{24}\right)}{\left(\frac{2}{3} - \left(-\frac{2}{3}\right)\right)} \\ &= \frac{5}{24} \div \frac{4}{3} \\ &= \frac{5}{32}. \end{aligned}$$

Probability Summary

If $a \le x_1 < x_2 \le b$ and we pick x at random between a and b, then:

$$P(x_1 < x < x_2) = \frac{\int_{x_1}^{x_2} w(x)\,dx}{\int_a^b w(x)\,dx} = \frac{\text{Part}}{\text{Whole}}.$$

In our previous example, the weighting function described the height of a curve above the x-axis.

Our next probability problem will be more realistic. Suppose you're throwing darts at a dart board and your little brother is standing next to the dart board. How likely are you to hit your little brother?

Errata: Heat is Energy

While computing the energy needed to boil the witch's cauldron last class, Professor Jerison said that we were computing *energy* and not *heat*. Energy, heat and work are all different names for the same thing, despite the fact that heat is measured in calories and work in foot-pounds. The only difference between the quantities is the units we use to describe them.

Example: Boy Near a Dart Board

Suppose a seven year old child is throwing darts at a dartboard while her little brother is standing nearby. What is the probability that the brother gets hit by a dart?

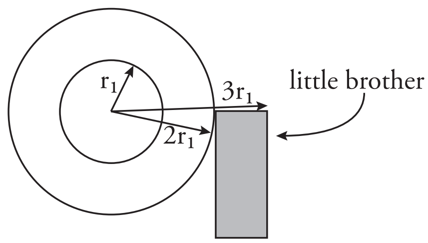


Figure 1: A younger brother stands by a dart board.

In order to turn this into a math problem we must make some assumptions. First, we'll assume something about her aim:

$$\text{Number of hits} = ce^{-r^2}$$

This says that the locations she hits are normally distributed (like a bell curve) about the center of the dart board. She's more likely to hit the dart board, but there's some chance she'll hit the wall next to it, or her brother.

As usual, we'll compute the probability by finding the ratio of the part to the whole. For the final calculation, the "part" will be where the brother is, and the "whole" is all the possible places the dart could hit.

We start by considering a thin ring or annulus of the dart board. The inside of the ring will have radius r_1 and the outside will have radius r_2.

If we graph $y = e^{-r^2}$ we get a "side view" of our probability distribution that looks like Figure 3. The height of the graph indicates that the darts are most likely to land near the center of the dartboard and less likely to land further out. (It turns out that the c in ce^{-r^2} will cancel, so we'll forget about it for now. The value of c depends on the number of darts thrown.)

To calculate the probability of a dart landing in the annulus described by r_1 and r_2 we look at the area between r_1 and r_2, above the r-axis, and below $y = e^{-r^2}$. We'll calculate the volume of revolution of this area, revolving about the center of the dart board (which corresponds to the y-axis in our graph). This volume of revolution will have a ring or washer shape.

Since we've written probability as a function of radius, we'll use the method of shells; the probability of a hit is constant at a given radius. To find the probability of hitting inside the annulus, we use the limits of integration r_1 and

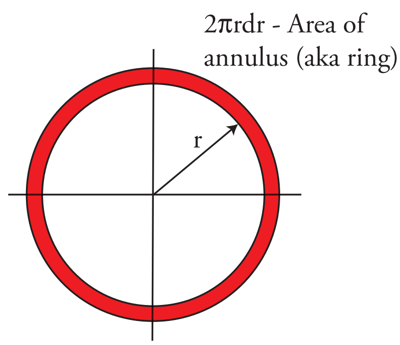


Figure 2: A annulus or ring about the center of the dartboard.

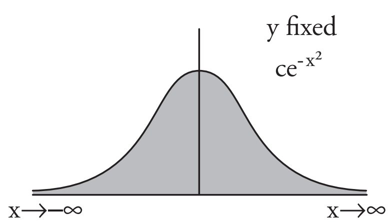


Figure 3: Graph of ce^{-x^2}.

r_2, circumference $2\pi r$, and height e^{-r^2}. We're integrating with respect to r so the thickness of a shell is dr.

$$\int_{r_1}^{r_2} (2\pi r)e^{-r^2}\,dr$$

Question: Why not just find the area under the curve between r_1 and r_2, then revolve that?

Answer: Because a shell with radius r_1 may have significantly less volume than one with radius r_2. We can't simply multiply that area by 2π to get an estimate of the volume of the ring.

Question: So r_1 and r_2 could be anything?

Answer: Yes. We'll do this in general first, and then later we can plug in values for r_1 and r_2.

Question: Why do we have to compute the volume?

Answer: We're computing the number of hits per unit area. The height of the graph corresponds to the number of hits, so we're computing area times height equals volume.

Question: Why is this a realistic calculation?

Answer: The function e^{-r^2} turns out to be the most accurate one for describing where the darts are likely to hit; this model was used to analyze where the V-2 rockets from Germany hit London during World War II.

We're constructing a mathematical model of dart throwing. The probability function e^{-r^2} models the fact that the kid is aiming for the center of the dartboard, but that due to inaccuracies in her throwing she won't always hit the bullseye. The assumption that the model describes the situation accurately does need to be justified, but for now Professor Jerison is asking us to accept it as-is.

Now we'll calculate the probability of a dart landing in the annulus between radii r_1 and r_2:

$$\begin{aligned}\int_{r_1}^{r_2} (2\pi r)e^{-r^2}\,dr &= -\pi e^{-r^2}\Big|_{r_1}^{r_2} \\ &= -\pi e^{-r_2^2} - (-\pi e^{-r_1^2}) \\ &= \pi(e^{-r_1^2} - e^{-r_2^2}).\end{aligned}$$

Note that the calculation doesn't change much if we instead use ce^{-r^2} as the probability function. We'll put it back now:

$$\text{Part} = c\pi(e^{-r_1^2} - e^{-r_2^2})$$

To calculate the volume of the "Whole", we let r range from zero to infinity. This is slightly artificial because when you're playing darts you don't hit the floor or the ceiling. But for the same reason, pretending the wall goes on forever doesn't have much effect on our final answer. The easiest value to calculate has the upper limit of integration equal to infinity; using this approximation will make the numbers come out more cleanly and because the e^{-r^2} is close to zero when r is large our answer will still be pretty accurate.

$$Whole = c\pi(e^{-0^2} - e^{-\infty^2}) = c\pi(1-0) = c\pi$$

The probability of hitting the annulus is then:

$$\begin{aligned}P(r_1 < r < r_2) &= \frac{\text{Part}}{\text{Whole}} \\ &= \frac{c\pi(e^{-r_1^2} - e^{-r_2^2})}{c\pi} \\ P(r_1 < r < r_2) &= e^{-r_1^2} - e^{-r_2^2}.\end{aligned}$$

(Notice that the c did, in fact, cancel.)

For this to be a realistic probability function, it should be true that

$$P(0 < r < \infty) = 1.$$

Since $e^0 = 1$ and $e^{-\infty^2} = 0$, this is in fact true. Another realistic assumption is that the probability of the child hitting the target is about 1/2. In other words, if a is the radius of the target,

$$P(0 \le r \le a) = \frac{1}{2}.$$

To figure out the probability of hitting the little brother, we have to define a "part" corresponding to the little brother. Let's suppose he's not standing too close — maybe he's standing a distance of $2a$ away from the dart board. Volumes of revolution are much easier to compute than volumes of preschoolers, so we'll assume that the parts of the little brother that are most likely to be hit lie on an arc of an annulus. We'll say that he's between $2a$ and $3a$ units away from the dartboard, and that he fills the angular arc between 3 o'clock and 5 o'clock to the right of the board, which is 1/6 of a circle.

So the probability of the little brother getting hit is:

$$\frac{1}{6}P(2a < r < 3a).$$

In order to get a number out of this we need to use the fact that the probability of the girl hitting the dart board is 1/2 or 50%.

$$\begin{aligned} P(0 < r < a) &= 1/2 \\ e^{0^2} - e^{-a^2} &= 1/2 \\ 1 - e^{-a^2} &= 1/2 \\ 1/2 &= e^{-a^2}. \end{aligned}$$

We could calculate a from this, but it turns out we don't need to. We now know all we need to calculate the probability of the little brother being hit. We start by calculating the probability of hitting the entire annulus between $2a$ and $3a$.

$$\begin{aligned} P(2a < r < 3a) &= e^{-(2a)^2} - e^{-(3a)^2} \\ &= e^{-2^2a^2} - e^{-3^2a^2} \\ &= e^{-4a^2} - e^{-9a^2} \\ &= (e^{-a^2})^4 - (e^{-a^2})^9 \\ &= \left(\frac{1}{2}\right)^4 - \left(\frac{1}{2}\right)^9 \\ &= \frac{1}{16} - \frac{1}{512} \\ &= \frac{1}{16} - \frac{1}{512} \\ &\approx \frac{1}{16}. \end{aligned}$$

(Getting the value $\frac{1}{512}$ for $r = 3a$ reassures us that the probably of getting hit does indeed decrease sharply as you get further from the dart board.)

So the probability that the girl hits her little brother is:

$$\begin{aligned}\frac{1}{6}P(2a < r < 3a) &\approx \frac{1}{6}\frac{1}{16} \\ &\approx \frac{1}{100} \\ &= 1\%.\end{aligned}$$

In this problem, our weight function was:

$$w(r) = 2\pi cre^{-r^2}.$$

This is different from the ce^{-r^2} that we started out with. That function is one dimensional; to take into account that we're describing the probability for a whole ring around the center of the dart board we include a factor of $2\pi r$.

If you graph this function you'll see that as r goes to 0 the value of $w(r)$ also goes to zero. This reflects the fact that it's hard to hit the center of the target because the center is very small. It's more likely that a dart will hit a ring around the center, because those rings have greater area than the bull's eye.

Introduction to Numerical Integration

Many functions don't have easy to describe antiderivatives, so many integrals must be (approximately) calculated by computer or calculator. These calculations also take the form of (simpler) weighted averages. There are many different techniques for computing numerical estimates of definite integrals. We'll go over three of these techniques.

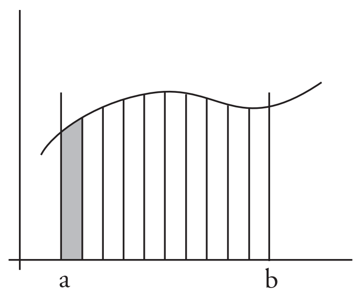


Figure 1: The area under the curve is divided into n regions of equal width.

- **Riemann Sums**

  Riemann sums are a very inefficient way to estimate the area under a continuous curve.

- **Trapezoidal Rule**

  Much more reasonable than Riemann sums, but still lousy.

- **Simpson's Rule**

  Slightly trickier, clever, and pretty good.

Review of Riemann Sums

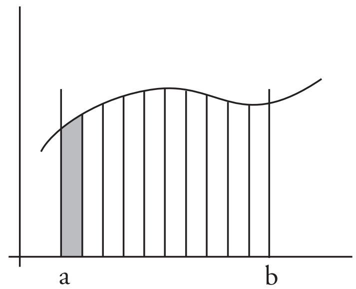


Figure 1: The area under the curve is divided into n regions of equal width.

As was mentioned at the start of this unit, Riemann sums approximate the area between the x-axis and a curve over the interval $[a, b]$ by a sum of areas of rectangles. Each rectangle has width $x_i - x_{i-1} = \Delta x$; there are n rectangles whose sides have x-coordinates $a = x_o < x_1 < x_2 ... < x_n = b$. The heights of the rectangles are $y_o = f(x_o),\ y_1 = f(x_1), ..., y_{n-1} = f(x_{n-1})$ (if the left edge of each rectangle is exactly as high as the graph).

Our goal is to "average" or add these y-values to get an approximation to

$$\int_a^b f(x)\, dx.$$

The formula for the (left) Riemann sum is:

$$(y_0 + y_1 + \ldots + y_{n-1})\Delta x.$$

If we let the right hand side of each rectangle be as high as the graph, using right endpoints instead of the left endpoints, we get the right Riemann sum:

$$(y_1 + y_2 + \ldots + y_n)\Delta x.$$

Trapezoidal Rule

$$\text{Area} \approx \Delta x \left(\frac{y_o}{2} + y_1 + y_2 + ... + y_{n-1} + \frac{y_n}{2} \right).$$

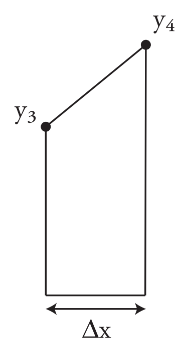


Figure 1: Approximation by areas of trapezoids.

The trapezoidal rule divides up the area under the graph into trapezoids (using segments of secant lines), rather than rectangles (using horizontal segments). As you can see from Figure 1, these diagonal lines come much closer to the curve than the tops of the rectangles used in the Riemann sum.

Remember that the area of a trapezoid is the area of the base times its average height. When applying the trapezoidal rule, the base of a trapezoid has length Δx and its sides have heights y_{i-1} and y_i; trapezoid i has area $\Delta x \frac{y_{i-1}+y+i}{2}$.

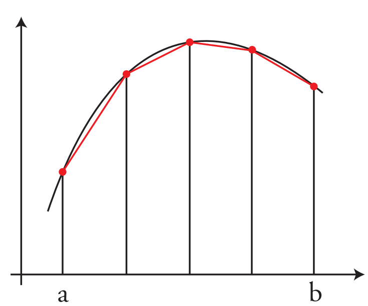


Figure 2: Area $= \left(\frac{y_3+y_4}{2}\right)\Delta x$.

When we add up the areas of all the trapezoids under the curve, we get:

$$\text{Area} \quad = \quad \Delta x \left\{ \frac{y_o + y_1}{2} + \frac{y_1 + y_2}{2} + \frac{y_2 + y_3}{2} + ... + \frac{y_{n-1} + y_n}{2} \right\}$$

$$= \quad \Delta x\left(\frac{y_o}{2} + y_1 + y_2 + ... + y_{n-1} + \frac{y_n}{2}\right).$$

Notice that the trapezoidal rule is the average of the left Riemann sum and the right Riemann sum; it gives a more symmetric treatment of the endpoints a and b than a Riemann sum does.

This looks good and in fact it is much better than a Riemann sum; however, it's still not very efficient.

Simpson's Rule

This approach often yields much more accurate results than the trapezoidal rule does. Again we divide the area under the curve into n equal parts, but for this rule n must be an even number because we're estimating the areas of regions of width $2\Delta x$.

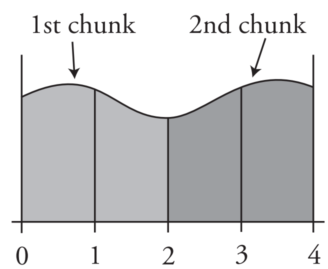


Figure 1: Simpson's rule for n intervals (n must be even!)

When computing Riemann sums, we approximated the height of the graph by a constant function. Using the trapezoidal rule we used a linear approximation to the graph. With Simpson's rule we match quadratics (i.e. parabolas), instead of straight or slanted lines, to the graph. When Δx is small this approximates the curve very closely, and we get a fantastic numerical approximation of the definite integral.

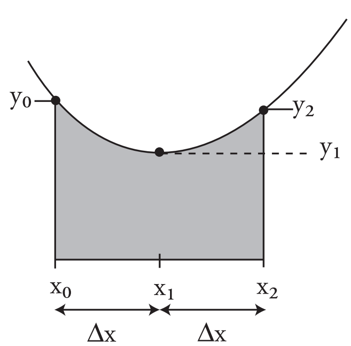


Figure 2: Using a parabolic approximation of the curve.

The derivation of the formula for Simpson's Rule is left as an exercise, but the area of this region is essentially the base times some average height of the graph:

$$\text{Area} = (\text{base})(\text{average height}) = (2\Delta x)\left(\frac{y_o + 4y_1 + y_2}{6}\right).$$

This emphasizes the middle more than the sides, which is consistent with the equations for parabolic approximation.

Simpson's rule gives you the following estimate for the area under the curve:

$$\text{Area} = (2\Delta x)\left(\frac{1}{6}\right)\left[(y_o + 4y_1 + y_2) + (y_2 + 4y_3 + y_4) + \cdots + (\cdots y_n)\right].$$

We can combine terms here by exploiting the following pattern in the coefficients:

1	4	1				
		1	4	1		
				1	4	1
1	4	2	4	1	4	1

To get the final form of Simpson's rule:

$$\int_a^b f(x)dx \approx \frac{\Delta x}{3}(y_o + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n).$$

Trapezoid Rule Approximation of $\int_1^2 \frac{dx}{x}$

Continuing our discussion of numerical integration, we'll look at:

$$\int_1^2 \frac{dx}{x}$$

Of course we already know:

$$\begin{aligned} \int_1^2 \frac{dx}{x} &= \ln x|_1^2 \\ &= \ln 2 - \ln 1 \\ &= \ln 2 \end{aligned}$$

We can use a calculator to find that this value is approximately 0.693147.

Numerical methods allow us to estimate integrals with accuracy about equal to our accuracy in estimating the integrand. We can approximate the value of $1/x$ pretty well, so we can get a pretty accurate estimate of the value of $\ln 2$. To make life even easier for ourselves, we'll do a very simple case of the approximation — we'll only use two intervals.

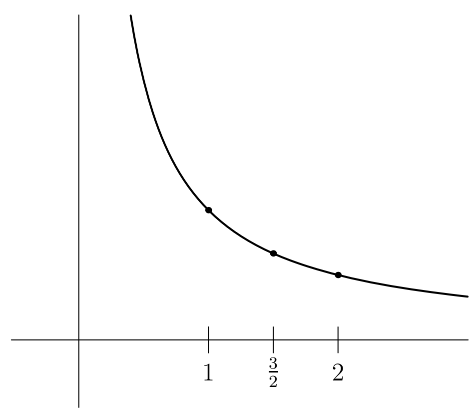


Figure 1: Two intervals; three points on the hyperbola.

We can't expect to get a very good approximation of $\ln 2$ using only two intervals. With only two intervals, we're making estimates of the area under a hyperbola based on only three points (see Figure 1).

Trapezoidal Rule

The trapezoidal rule gives us the following formula for the area under the curve:

$$\text{Area} \approx \Delta x \left(\frac{1}{2} y_0 + y_1 + \frac{1}{2} y_2 \right).$$

In this case $\Delta x = \frac{b-a}{n} = \frac{1}{2}$ because $b = 2$, $a = 1$ and $n = 2$. By evaluating $y_i = f(x_i) = \dfrac{1}{1+\frac{1}{2}i}$ and plugging these values in, we get:

$$\text{Area} \approx \frac{1}{2}\left(\frac{1}{2}\cdot 1 + \frac{2}{3} + \frac{1}{2}\cdot\frac{1}{2}\right)$$

You wouldn't have to perform this addition on an exam, but if there are only two terms you should add them together. When we do put these numbers into a calculator, we find that the trapezoidal rule gives an estimate of about 0.708 for the area of this region; that's pretty close under the circumstances!

Simpson's Rule Approximation of $\int_1^2 \frac{dx}{x}$

If we use Simpson's rule with two intervals to estimate the value of $\int_1^2 \frac{dx}{x}$, the result is surprisingly accurate:

$$\frac{\Delta x}{3}(y_0 + 4y_1 + y_2) = \frac{1}{6}\left(1 + 4\frac{2}{3} + \frac{1}{2}\right) \approx 0.69444...$$

This is impressively close to our calculator's estimate of 0.693147. (Our calculator uses more than two intervals!)

The accuracy of Simpson's rule is proportional to Δx^4. In this example, that means that if we'd used 10 intervals the error would be about 10^{-4}; we'd have four digit accuracy. The calculation isn't even particularly difficult — it's possible to do it by hand. The other rules for numerical approximation aren't as accurate.

The reason Simpson's rule is more accurate is that it's matching a parabola to the curve, rather than a straight line. Simpson's rule gives the exact area beneath the graphs of functions of degree two or less (parabolas and straight lines), while the other methods are only exact for functions whose graphs are linear.

Oddly, Simpson's rule also gives the exact area under the curve for cubic curves, which explains why the error is comparable to the fourth power of Δx.

However, Simpson's rule does have problems if the derivative is unbounded (e.g. near 0 for the function $f(x) = \frac{1}{x}$) or if the graph is not "smooth" enough.

Numerical Integration Study Tips

How can we remember the formulas for the trapezoidal rule and Simpson's rule to use on the exam? How can we know if we've remembered the formulas correctly?

Instead of memorizing the entire formula, you might memorize a very simple case:

$$f(x) = 1$$

and reconstitute the complete formula from that. The formula for $f(x) = 1$ is easy to check — if the area under the curve doesn't come out to $b - a$, you're doing it wrong.

Consider the formula for the trapezoid rule:

$$\text{Area} \approx \Delta x \left(\frac{1}{2} y_0 + y_1 + \cdots + y_{n-1} + \frac{1}{2} y_n \right).$$

If the function is $f(x) = 1$, then $y_i = 1$ for each value of i:

$$\begin{aligned}
\Delta x \left(\frac{1}{2} + \overbrace{1 + 1 + \cdots + 1}^{(n-1) \text{ of these}} + \frac{1}{2} \right) &= \Delta x \left(\frac{1}{2} + (n-1) + \frac{1}{2} \right) \\
&= \Delta x \cdot n \\
&= \frac{b-a}{n} \cdot n \\
&= b - a \\
&= \int_a^b 1 \, dx.
\end{aligned}$$

You can do the same thing with Simpson's rule.

Area Under the Bell Curve

Today, we'll complete the calculation first mentioned during the discussion of the fundamental theorem of calculus. We've said that:

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}.$$

Now we'll prove it. The proof relies on a very clever trick which we would be unlikely to come up with ourselves. We study the proof because the result is very important and because it's related to adding up slices, as we've been doing in this unit.

In our example with the little brother and the dart board, we found the volume of revolution created by rotating the curve e^{-r^2} around the vertical axis.

The calculation went by shells:

$$V = \int_0^\infty 2\pi r\ e^{-r^2}\,dr$$

where $2\pi r$ was the circumference of the shell, e^{-r^2} was the height, and dr was the thickness.

$$\begin{aligned} V &= \int_0^\infty 2\pi r\, e^{-r^2}\,dr \\ &= -\pi\ e^{-r^2}\Big|_0^\infty \\ V &= \pi \end{aligned}$$

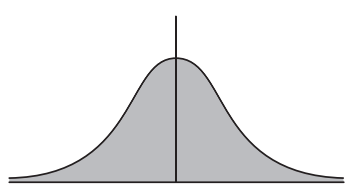


Figure 1: $Q =$ area under e^{-t^2}.

What we need to find now is not a volume but an area:

$$Q = \int_{-\infty}^\infty e^{-t^2}\,dt$$

This is the area under the bell curve shown in Figure 1.

The trick is to compute V in a different way — by slices. Amazingly, we'll discover that $V = Q^2$, which will tell us the value of Q:

$$\begin{aligned} Q^2 &= V \\ Q^2 &= \pi \\ Q &= \sqrt{\pi}. \end{aligned}$$

So now we need to check that $V = Q^2$.

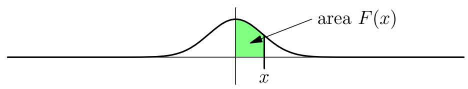


Figure 2: $F(x) = \int_0^x e^{-t^2 dt}\, dt$.

Remember that we have already discussed:

$$F(x) = \int_0^x e^{-t^2}\, dt.$$

We were interested in the limit of $F(x)$ as x approached infinity:

$$F(\infty) = \int_0^\infty e^{-t^2}\, dt.$$

That's the area under half of the bell curve, so:

$$\begin{aligned} Q &= 2F(\infty) \\ F(\infty) &= \frac{\sqrt{\pi}}{2} \end{aligned}$$

By calculating the value of Q we're reassuring ourselves that this is true.

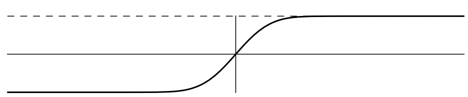


Figure 3: $\lim_{x\to\infty} F(x) = \frac{\sqrt{\pi}}{2}$

We need to prove $V = Q^2$ by using slices; two slices of the surface e^{-r^2} are shown in Figure 4. This is not particularly easy to visualize. It may help to imagine slicing an anthill, a pile of gravel, or some other bump that has circular symmetry.

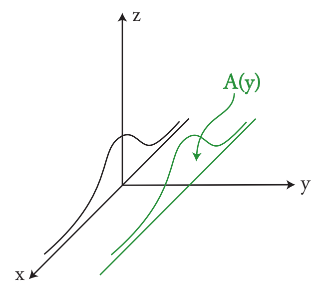


Figure 4: Three dimensional slices of the volume of rotation of e^{-r^2}.

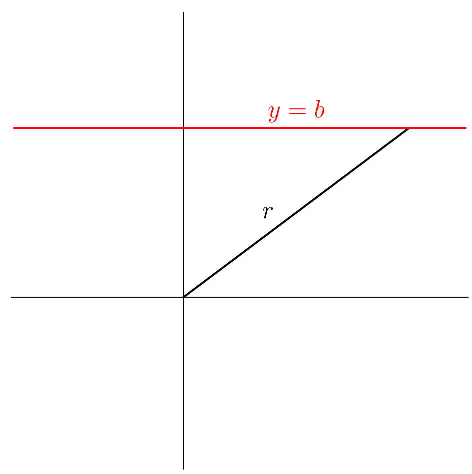


Figure 5: Top view of a slice of the surface of revolution of e^{-r^2}.

The formula for volume by slices is:

$$V = \int_{-\infty}^{\infty} A(y)\, dy.$$

We're going to fix $y = b$ and calculate $A(b)$.

Figure 5 shows a top view of the slice whose area we're calculating. The height of the surface at a point r units away from $(0, 0)$ is given by:

$$\text{height} = e^{-r^2}.$$

In terms of b and x, $r^2 = b^2 + x^2$, so

$$\begin{aligned} \text{height} &= e^{-(b^2+x^2)} \\ &= e^{-b^2} e^{-x^2} \\ &= ce^{-x^2} \end{aligned}$$

where c is a constant equal to e^{-b^2}.

The area under the curve is:

$$\begin{aligned} A(b) &= \int_{\infty}^{\infty} e^{-b^2} e^{-x^2}\, dx \\ &= e^{-b^2} \int_{\infty}^{\infty} e^{-x^2}\, dx \\ A(b) &= e^{-b^2} Q \end{aligned}$$

We've calculated the area of a slice of our "bump" in terms of the value Q that we're looking for.

Remember that the volume of that bump is given by:

$$\begin{aligned} V &= \int_{-\infty}^{\infty} A(y)\, dy \\ &= \int_{-\infty}^{\infty} e^{-y^2} Q\, dy \\ &= Q \int_{-\infty}^{\infty} e^{-y^2}\, dy \quad (Q \text{ is a constant}) \\ &= Q^2 \quad (\text{by definition, } Q = \textstyle\int_{-\infty}^{\infty} e^{-t^2}\, dt). \end{aligned}$$

As desired, we've proven that $V = Q^2$ and so we can conclude that $Q = \sqrt{\pi}$.

Question: Doesn't x change as y changes?

Answer: Good question. That's the way x and y have been used in this whole course. But x and y can be different variables; they won't always depend on each other.

In this example, we're fixing $y = b$. The value of y doesn't change. The value of x does change; x varies from negative infinity to positive infinity. In this example y is not a function of x, and that's ok but it's not what we're used to.

Questions on Test 3

The Unit 3 Exam will consist of five questions. It will be very similar to the Fall 2005 exam.

1. Calculate Definite Integrals (via FTC and Substitution)
2. Numerical Approximation:
   Riemann Sum
   Trapezoidal Rule
   Simpson's Rule
3. Areas/Volumes
4. Other Cumulative Sums (Average Value, Probability, Work)
5. Sketch the Graph of $F(x) = \int_a^x f(t)\,dt$.

Types of Riemann Sums

Question: For Riemann sums, what's the difference between upper and lower, and right and left?

Answer: If you take a function like $f(x) = \frac{1}{x}$ and break it up into pieces, the lower sum is the sum of the areas of rectangles which are always *lower* than the graph of the function; see Figure 1. The right sum is the sum of the areas

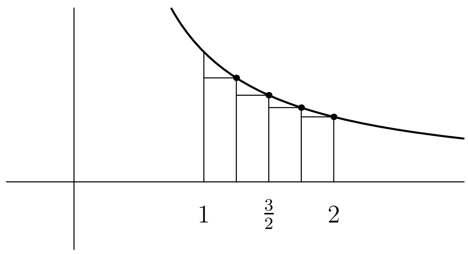


Figure 1: Lower sum.

of rectangles whose heights are $f\left(\frac{(i+1)\Delta x}{n}\right)$. If the values of f are positive, the upper right corner of each rectangle lies on the graph of f.

For a continuously decreasing function like $\frac{1}{x}$, the lower sum equals the right sum and the upper sum equals the left sum.

If the graph of the function "wiggles" up and down, the upper sum will be the sum of the areas of rectangles whose height is the maximum value the function achieves on the interval $\left[\frac{i\Delta x}{n}, \frac{(i+1)\Delta x}{n}\right]$. The maximum could be achieved at the left endpoint, the right endpoint or somewhere in the middle of the interval.

Question: If the function is increasing, then the lower sum is the left sum?

Answer: Correct. It's the exact opposite of the situation with an increasing function.

Asymptotes of Antiderivatives

Question: Suppose that the function $F(x) = \int_a^x f(t)\,dt$ in question 5 has a horizontal asymptote. How do we tell where, or even if, there's an asymptote?

Answer: The questions about $F(x) = \int_a^x f(t)\,dt$ are generally the trickiest problems. During the exam you probably won't be able to figure out what happens to the function as x goes to infinity. Just deciding whether there is an asymptote or not is a serious question which we'll address at the very end of the course.

Choosing a Technique

Question: How will we know which method to use on the exam?

Answer: You'll need all three methods of Riemann sums, trapezoidal rule and Simpson's rule.

Volumes of revolution always depend on a two dimensional diagram which you should be able to figure out. (There won't be any volumes as complicated as the surface e^{-r^2} on the exam.)

Once you've got the two dimensional diagram for a volume or area calculation, you'll have a choice between integrating with respect to x or y. If there's some choice that seems reasonable but makes the problem very hard, you might get a hint suggesting you use the washer method, for example. If, on the other hand, there's one choice that's reasonable and another that's obviously very difficult, you may be left to figure out which is the right choice on your own.

For example, suppose you're looking at the volume formed by rotating the region $0 < y < x - x^3$ shown in Figure 1 about the y-axis.

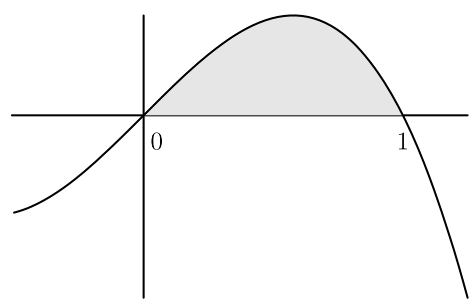


Figure 1: $0 < y < x^3 - x$.

We have to choose whether to integrate with respect to x or with respect to y. If we choose to integrate with respect to x, we'll be adding up volumes of cylindrical shells with thickness dx and height $x - x^3$.

$$\int_0^1 2\pi x(x - x^3)\, dx$$

When we reach this point, we know we've made a good choice because this is an easy integral to calculate.

If, on the other hand, we decided to integrate with respect to y, we'd be adding up volumes of "washers" – disks with circular holes in their centers. (See Figure 2.) Using the washer method, we get the following formula for the volume:

$$\int \pi(x_2^2 - x_1^2)\, dy$$

We can already see that this is more complicated than the other, but let's discuss another step of the calculation to see how much more complicated. To compute this integral, we'll need to solve $y = x - x^3$ for x in terms of y. It's not easy to

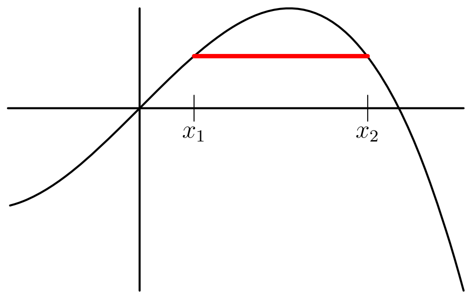


Figure 2: Rotate this rectangle about the y-axis to get a washer shaped volume.

solve $x^3 - x + y = 0$; you won't be able to do it during the test. So integrating with respect to y is the wrong choice for this problem.

Review of Trigonometric Identities

The topic of this segment is the use of trigonometric substitutions in integration. We start by reviewing some basic facts about trigonometry.

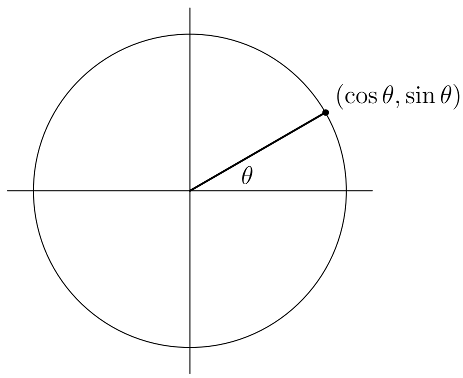


Figure 1: The unit circle.

Trigonometry is based on the circle of radius 1 centered at $(0, 0)$. A point on that circle at angle θ (see Figure **??**fig:l27g1) has coordinates $(\cos\theta, \sin\theta)$. Because the radius of the circle is 1, the Pythagorean theorem tells us right away that $\sin^2\theta + \cos^2\theta = 1$. (Remember that $\sin^2\theta$ means $(\sin\theta)^2$.) You may also remember some double angle formulas.

$$\begin{aligned}
\sin^2\theta + \cos^2\theta &= 1 \\
\cos(2\theta) &= \cos^2\theta - \sin^2\theta \\
\sin(2\theta) &= 2\sin\theta\cos\theta
\end{aligned}$$

From the double angle formula for $\cos(2\theta)$ we can derive the half angle formula:

$$\begin{aligned}
\cos(2\theta) &= \cos^2\theta - \sin^2\theta \\
&= \cos^2\theta - (1 - \cos^2\theta) \\
\cos(2\theta) &= 2\cos^2\theta - 1 \\
\Rightarrow \cos^2\theta &= \frac{1 + \cos(2\theta)}{2}
\end{aligned}$$

This formula will allow us to rewrite powers like $\cos^2\theta$ in lower degree terms. A similar calculation shows that:

$$\sin^2\theta = \frac{1 - \cos(2\theta)}{2}.$$

$\int \sin^n x \cos^m x \, dx, m = 1$

You already know something about integrating trigonometric functions; you can reverse anything you know about the derivative of a trigonometric function to get a fact about antiderivatives.

$$\begin{aligned} d\sin x = \cos x \, dx \quad &\Rightarrow \quad \int \cos x \, dx = \sin x + c \\ d\cos x = -\sin x \, dx \quad &\Rightarrow \quad \int \sin x \, dx = -\cos x + c \end{aligned}$$

Our plan is to use these two integration formulas and a few trig identities to derive more complicated formulas involving trig functions.

Our first topic is integrals of the form:

$$\int \sin^n x \cos^m x \, dx,$$

where m and n are non-negative integers. Integrals like this appear in Fourier series, among other places.

There are two cases to think about here. The easy case is the one in which at least one exponent is odd.

Example: $m = 1$

The trick in calculating $\int sin^n(x)\cos(x)\,dx$ is to make the substitution $u = \sin x$, so $du = \cos x \, dx$.

$$\begin{aligned} \int sin^n(x)\cos(x)\,dx &= \int u^n du \\ &= \frac{u^{n+1}}{n+1} + c \\ &= \frac{\sin^{n+1} x}{n+1} + c \end{aligned}$$

Although the answer $\frac{u^{n+1}}{n+1} + c$ looks nice, you need to reverse your substitution and plug in $\sin x$ for u at the end because while you may know what u is, your grader or employer does not.

What made this problem easy is that $\cos x$ is the derivative of $\sin x$.

Example: $\int \sin^3 x \cos^2 x dx$

One of the exponents is odd so this is an easy, but not as easy as the previous example. We turn this integral into one in which the odd exponent is 1 by using the trig identity $\sin^x x + \cos^2 x = 1$ to remove the largest even power in the term with the odd exponent.

In this case the odd exponent is on $\sin x$, so we use:

$$\sin^2 x = 1 - \cos^2 x.$$

$$\begin{aligned}
\int \sin^3 x \cos^2 x dx &= \int (\sin^2 x \cdot \sin x) \cos^2 x dx \\
&= \int (1 - \cos^2 x) \cdot \sin x \cos^2 x dx \\
&= \int (\cos^2 - \cos^4 x) \sin x dx
\end{aligned}$$

This is now very similar to our previous example. We use the substitution:

$$\begin{aligned}
u &= \cos x \\
du &= -\sin x dx
\end{aligned}$$

To get:

$$\begin{aligned}
\int \sin^3 x \cos^2 x &= \int (u^2 - u^4 x) \cdot (-du) \\
&= \int \frac{-u^3}{3} + \frac{u^6}{5} + c \\
&= -\frac{\cos^3 x}{3} + \frac{\cos^5 x}{5} + c
\end{aligned}$$

At this point it's a good idea to check your work by differentiating your answer and then applying trig identities to see that the result equals the original integrand.

Example: $\int \sin^3 x dx$

The integral $\int \sin^3 x dx$ is of the form $\int \sin^n x \cos^m x dx$ with one exponent odd, and the other exponent equal to zero, so it is in the easy case. We again use the trig identity $\sin^2 x + \cos^2 x = 1$ to remove the largest power of $\sin x$ that we can from the cube:

$$\int \sin^3 x dx = \int (1 - \cos^2 x) \sin x dx$$

Substitute $u = \cos x$ and $du = -\sin x dx$ to get:

$$\begin{aligned} \int \sin^3 x dx &= \int (1 - u^2)(-du) \\ &= -u + \frac{u^3}{3} + c \\ &= -\cos x + \frac{\cos^3 x}{3} + c \end{aligned}$$

In general, any time you have an odd power in an integral of the form $\int \sin^n x \cos^m x dx$ you can integrate it using the trig identity $\sin^2 x + \cos^2 x = 1$ and a substitution.

Example: $\int \cos^2 x\, dx$

What if we have to integrate $\int sin^n x \cos^m x\, dx$ when both exponents are even? This is a harder case; we'll use the half angle formulas to solve it.

$$\cos^2\theta = \frac{1+\cos(2\theta)}{2}$$
$$\sin^2\theta = \frac{1-\cos(2\theta)}{2}$$

These formulas help us by turning even powers of $\sin x$ and $\cos x$ into odd powers of $\cos(2x)$.

If we wanted to integrate:

$$\int \cos^2 x\, dx,$$

we could rewrite it as $\int (1-\sin^2 x)\, dx$, but the new integral is at least as hard as the one we started with. Instead we use a half angle formula:

$$\begin{aligned} \int \cos^2 x\, dx &= \int \frac{1+\cos(2x)}{2}\, dx \\ &= \frac{x}{2} + \frac{\sin 2x}{4} + c \end{aligned}$$

Notice that $\frac{x}{2}$ appears in the solution and is not a trigonometric function!

Example: $\displaystyle\int \sin^2 x \cos^2 x \, dx$

To integrate $\sin^2 x \cos^2 x$ we once again use the half angle formulas:

$$\cos^2\theta = \frac{1+\cos(2\theta)}{2}$$
$$\sin^2\theta = \frac{1-\cos(2\theta)}{2}$$

It's a good idea to do your trigonometric and algebraic manipulations of the integrand off to the side on your paper, so that you don't have to continuously copy over (and maybe forget) the integral sign and the dx.

Side work:

$$\begin{aligned} \sin^2 x \cos^2 x &= \left(\frac{1-\cos(2x)}{2}\right)\left(\frac{1+\cos(2x)}{2}\right) \\ &= \frac{1-\cos^2(2x)}{4} \end{aligned}$$

We still have a square, so we're still not in the easy case. But this is an easier "hard" case, especially since we just computed $\int \cos^2 x \, dx$. We could use that, but instead let's continue to use half angle formulas until we reach an easy case:

$$\begin{aligned} \sin^2 x \cos^2 x &= \frac{1-\cos^2(2x)}{4} \\ &= \frac{1}{4} - \frac{1+\cos(4x)}{4\cdot 2} \\ &= \frac{1}{8} - \frac{\cos(4x)}{8} \end{aligned}$$

Once we've done the side work we substitute back into the original integral to get:

$$\begin{aligned} \int \sin^2 x \cos^2 x \, dx &= \int \left(\frac{1}{8} - \frac{\cos(4x)}{8}\right) dx \\ &= \frac{x}{8} - \frac{\sin(4x)}{8\cdot 4} + c \end{aligned}$$

We should now be able to calculate any integral of the form $\int \sin^n x \cos^m x \, dx$.

Here's an alternate method of doing the side work using the identity

$$\sin(2\theta) = 2\sin\theta\cos\theta.$$

$$
\begin{aligned}
\sin^2 x \cos^2 x &= (\sin x \cos x)^2 \\
&= \left(\frac{1}{2}\sin(2x)\right)^2 \\
&= \frac{1}{4}\sin^2(2x) \\
&= \frac{1}{4}\left(\frac{1-\cos(4x)}{2}\right) \\
\sin^2 x \cos^2 x &= \frac{1}{8} - \frac{\cos(4x)}{8}
\end{aligned}
$$

Using the double angle formula for the sine function reduces the number of factors of $\sin x$ and $\cos x$, but not quite far enough; it leaves us with a factor of $\sin^2(2x)$. Next, the half angle formula for the sine function allows us to reduce this to a constant minus a multiple of the cosine function.

Note that we get the same expression we did before.

Area of Part of a Circle

Given a circle of radius a, cut out a tab of height b. What is the area of this tab? (See Figure 1.)

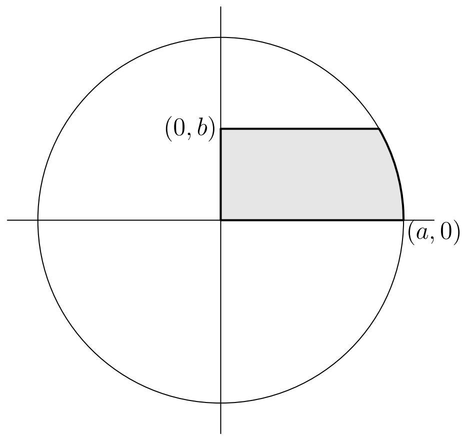


Figure 1: Tab cut out of a circle.

One way to compute the area would be split the area into vertical strips and integrate with respect to x:

$$\text{Area} = \int y\,dx.$$

This is awkward, because near the end the height of the region changes from a constant $y = b$ to the height of the circle $y = \sqrt{a^2 - x^2}$.

What if we integrate with respect to y? That seems to work better; there is a single simple expression for the length of each horizontal strip: $x = \sqrt{a^2 - y^2}$.

$$\begin{aligned} \text{Area} &= \int_0^b x\,dy \\ &= \int_0^b \sqrt{a^2 - y^2}\,dy \end{aligned}$$

We don't yet have a rule for integrating functions of this form. Considering that this integral arose from a question about a circle, it's not surprising that trigonometry will play a role in its solution.

When working with circles it often helps to use polar coordinates. In this case, note that the upper right hand corner of the region has polar coordinates $(a\cos\theta_0, a\sin\theta_0)$ where θ_0 is the angle shown in Figure 2.

In general, $x = a\cos\theta$ and $y = a\sin\theta$. If we substitute $y = a\sin\theta$ into our integrand we get:

$$x = \sqrt{a^2 - y^2}$$

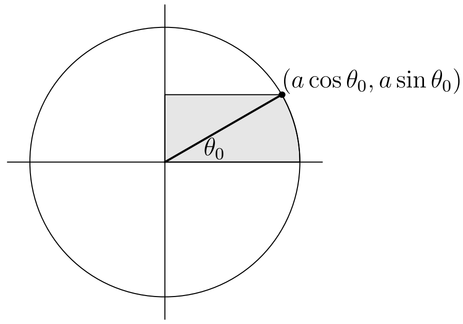


Figure 2: Polar coordinates of a point.

$$\begin{aligned} &= \sqrt{a^2 - a^2\sin^2\theta} \\ &= a\sqrt{1-\sin^2\theta} \\ &= a\sqrt{\cos^2\theta} \\ x &= a\cos\theta \end{aligned}$$

Changing to polar coordinates made our integrand look much nicer; we've gone from an integrand with a square root and no trig functions to an integrand with trig functions and no square root.

If we're going to use the substitution $y = a\sin\theta$ in our integral, we'll also need to replace dy by something in polar coordinates.

$$\begin{aligned} y &= a\sin\theta \\ dy &= a\cos\theta\, d\theta \end{aligned}$$

Plugging in, we get:

$$\begin{aligned} \int \sqrt{a^2-y^2}\, dy &= \int (a\cos\theta)(a\cos\theta\, d\theta) \\ &= a^2 \int \cos^2\theta\, d\theta \end{aligned}$$

We computed the integral of $\cos^2 x$ earlier in the lecture:

$$\int \cos^2 x\, dx = \frac{x}{2} + \frac{\sin 2x}{4} + C.$$

Plugging this in, we get:

$$\int \sqrt{a^2-y^2} dy = a^2\left(\frac{\theta}{2} + \frac{\sin(2\theta)}{4}\right) + c.$$

Now we'd like to rewrite our solution in terms of the original variable y so that we can plug in the limits of integration. In order to do this, it's helpful to rewrite $\sin(2\theta)$ using the double angle formula $\sin(2\theta) = 2\sin\theta\cos\theta$.

$$\begin{aligned}
\int \sqrt{a^2-y^2}\,dy &= a^2\left(\frac{\theta}{2}+\frac{\sin(2\theta)}{4}\right)+c \\
&= a^2\left(\frac{\theta}{2}+\frac{\sin\theta\cos\theta}{2}\right)+c \\
&= \left(\frac{a^2\theta}{2}+\frac{a\sin\theta a\cos\theta}{2}\right)+c.
\end{aligned}$$

Next we solve $y = a\sin\theta$ for y and plug in:

$$\theta = \arcsin\left(\frac{y}{a}\right).$$

Since $a\sin\theta = y$ and $a\cos\theta = x = \sqrt{a^2-y^2}$, we get:

$$\int \sqrt{a^2-y^2}\,dy = \left(\frac{a^2\arcsin(y/a)}{2}+\frac{y\sqrt{a^2-y^2}}{2}\right)+c.$$

We've used trigonometric substitution to find the indefinite integral of $\sqrt{a^2-y^2}$. Whenever you see the square root of a quadratic in an integral you should think of trigonometry and $\sin^2\theta+\cos^2\theta$.

Our original problem asked us to compute the value of a definite integral; let's finish that.

$$\begin{aligned}
\int_0^b \sqrt{a^2-y^2}\,dy &= \left.\left(\frac{a^2\arcsin(y/a)}{2}+\frac{y\sqrt{a^2-y^2}}{2}\right)\right|_0^b \\
&= \left(\frac{a^2\arcsin(b/a)}{2}+\frac{b\sqrt{a^2-b^2}}{2}\right)-0 \\
&= \frac{a^2\arcsin(b/a)}{2}+\frac{b\sqrt{a^2-b^2}}{2}
\end{aligned}$$

Notice that $\theta_0 = \arcsin(b/a)$; we could rewrite this answer as:

$$\text{Area} = \frac{a^2\theta_0}{2}+\frac{b\sqrt{a^2-b^2}}{2}$$

Does this make sense? The first term, $\frac{\theta_0}{2}a^2$, is exactly the area of the sector of the circle swept out by angle θ_0. The second term, $\frac{1}{2}b\sqrt{a^2-b^2}$, is the area of a triangle with base b and height $\sqrt{a^2-b^2}$. In other words, it's the area of the shaded triangle shown in Figure 2.

Using some basic geometry, we've checked that our answer to this complicated calculus problem is correct.

Review of Trigonometric Identities

We've talked about trig integrals involving the sine and cosine functions. Now we'll look at trig functions like secant and tangent. Here's a quick review of their definitions:

$$\sec x = \frac{1}{\cos x} \qquad \tan x = \frac{\sin x}{\cos x} \tag{1}$$

$$\tag{2}$$

$$\csc x = \frac{1}{\sin x} \qquad \cot x = \frac{\cos x}{\sin x} \tag{3}$$

When you put a "co" in front of the name of the function, that exchanges the roles of sine and cosine in that function.

We have the following identities:

$$\begin{aligned} \sec^2 x &= 1 + \tan^2 x \\ \frac{d}{dx} \tan x &= \sec^2 x \\ \frac{d}{dx} \sec x &= \sec x \tan x \end{aligned}$$

We can verify these using familiar trig identities involving $\sin x$ and $\cos x$.

$$\begin{aligned} \sec^2 x &= \frac{1}{\cos^2 x} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \\ \sec^2 x &= 1 + \tan^2 x \end{aligned}$$

This is the main trig identity behind what we'll do today.

$$\begin{aligned} \frac{d}{dx} \tan x &= \frac{d}{dx}\left(\frac{\sin x}{\cos x}\right) \\ &= \frac{(\cos x)(\cos x) - (\sin x)(-\sin x)}{\cos^2 x} \quad \text{(chain rule)} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \\ &= \frac{1}{\cos^2 x} \\ \frac{d}{dx} \tan x &= \sec^2 x \end{aligned}$$

From this we get our first integral of the day:

$$\int \sec^2 x \, dx = \tan x + c.$$

$$
\begin{aligned}
\frac{d}{dx}\sec x &= \frac{d}{dx}\frac{1}{\cos x} \\
&= \frac{0-(-\sin x)}{\cos^2 x} \\
&= \frac{\sin x}{\cos^2 x} \\
\frac{d}{dx}\sec x &= \tan x \sec x
\end{aligned}
$$

Should we ever need an antiderivative of $\tan x \sec x$ we now have one.

Integral of Tangent

How do we integrate one of these trig functions if we can't work backward from a derivative we already know?

$$\int \tan x \, dx \quad = \quad \int \frac{\sin x}{\cos x} \, dx$$

If you're working an integral like this and you see a trig function, it's good to look around and see if you can also find the derivative of that trig function. We make the substitution:

$$u = \cos x, \quad du = -\sin x \, dx$$

and rewrite our integral as:

$$\begin{aligned}
\int \tan x \, dx &= \int \frac{\sin x \, dx}{\cos x} \\
&= \int \frac{-du}{u} \\
&= -\ln |u| + c \\
\int \tan x \, dx &= -\ln |\cos(x)| + c
\end{aligned}$$

You'll find tables of formulas like this in the back of most textbooks. In addition, there is a certain amount of memorization that goes on in calculus; this is the kind of thing that you probably want to memorize.

Integral of Secant

$$\int \sec x \, dx = ?$$

This calculation is not as straightforward as the one for the tangent function. What we need to do is add together the formulas for the derivatives of the secant and tangent functions.

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

Notice that $\sec x + \tan x$ appears on both sides of the equation here. If we let $u = \sec x + \tan x$ and substitute, our equation becomes:

$$u' = u \cdot \sec x.$$

Which tells us that:

$$\sec x = \frac{u'}{u}.$$

We've seen this before; this is called the *logarithmic derivative*:

$$\frac{u'}{u} = \ln(u).$$

Putting this all together in order, we get:

$$\begin{aligned} \sec x &= \frac{u'}{u} \quad (u = \sec x + \tan x) \\ &= \frac{d}{dx} \ln u \\ \sec x &= \frac{d}{dx} \ln(\sec x + \tan x). \end{aligned}$$

Integrating both sides, we get:

$$\int \sec x \, dx = \ln(\sec x + \tan x) + c.$$

By taking the derivative of exactly the right function and looking at the results in the right way we got the formula we needed. You won't be expected to do this yourself in this class.

Summary of Trig Integration

We now know the following facts about trig functions and calculus:

$$\sec x = \frac{1}{\cos x} \qquad \tan x = \frac{\sin x}{\cos x} \qquad \sin^2 x + \cos^2 x = 1$$

$$\csc x = \frac{1}{\sin x} \qquad \cot x = \frac{\cos x}{\sin x} \qquad \sec^2 x = 1 + \tan^2 x$$

$$\tfrac{d}{dx} \tan x = \sec^2 x \qquad \tfrac{d}{dx} \sec x = \sec x \tan x \qquad \textstyle\int \tan x \, dx = -\ln|\cos(x)| + c$$

$$\tfrac{d}{dx} \sin x = \cos x \qquad \tfrac{d}{dx} \cos x = -\sin x \qquad \textstyle\int \sec x \, dx = \ln(\sec x + \tan x) + c$$

We've also seen several useful integration techniques, including methods for integrating any function of the form $\sin^n x \cos^m x$. At this point we have the tools needed to integrate most trigonometric polynomials.

Example: $\displaystyle\int \sec^4 x \, dx$

We can get rid of some factors of $\sec x$ using the identity $\sec^2 x = 1 + \tan^2 x$. This is a particularly good idea because $\sec^2 x$ is the derivative of $\tan x$.

$$\begin{aligned} \int \sec^4 x \, dx &= \int (1 + \tan^2 x) \sec^2 x \, dx \\ &= \int (1 + \tan^2 x) \sec^2 x \, dx. \end{aligned}$$

Using the substitution $u = \tan x$, $du = \sec^2 x \, dx$, we get:

$$\begin{aligned} \int \sec^4 x \, dx &= \int (1 + u^2) du \\ &= u + \frac{u^3}{3} + c \\ \int \sec^4 x \, dx &= \tan x + \frac{\tan^3 x}{3} + c. \end{aligned}$$

Example of Trig Substitution: $\displaystyle\int \frac{dx}{x^2\sqrt{1+x^2}}$

$$\int \frac{dx}{x^2\sqrt{1+x^2}} = ?$$

This is an ugly integral. The square root is the ugliest part, so we'll try to rewrite it in such a way that we can get rid of the square. If we let $x = \tan\theta$ then the identity $\sec^2\theta = 1 + \tan^2\theta$ will allow this. We'll then have $dx = \sec^2\theta\, d\theta$:

$$\begin{aligned}
\int \frac{dx}{x^2\sqrt{1+x^2}} &= \int \frac{\sec^2\theta\, d\theta}{\tan^2\theta\sqrt{1+\tan^2\theta}} \\
&= \int \frac{\sec^2\theta\, d\theta}{\tan^2\theta\sqrt{\sec^2\theta}} \\
&= \int \frac{\sec^2\theta\, d\theta}{\tan^2\theta\sec\theta} \\
&= \int \frac{\sec\theta\, d\theta}{\tan^2\theta}
\end{aligned}$$

When faced with an assortment of different trig functions like this one, it's a good idea to rewrite everything in terms of $\sin\theta$ and $\cos\theta$:

$$\begin{aligned}
\int \frac{dx}{x^2\sqrt{1+x^2}} &= \int \frac{\frac{1}{\cos\theta}\, d\theta}{\frac{\sin^2\theta}{\cos^2\theta}} \\
&= \int \frac{\cos^2\theta\, d\theta}{\cos\theta\sin^2\theta} \\
&= \int \frac{\cos\theta\, d\theta}{\sin^2\theta}
\end{aligned}$$

The ugliest part of this integral is the $\sin^2\theta$ in the denominator. Since $\cos\theta\, d\theta$ is the derivative of $\sin\theta$, we make the substitution $u = \sin\theta$, $du = \cos\theta\, d\theta$:

$$\begin{aligned}
\int \frac{dx}{x^2\sqrt{1+x^2}} &= \int \frac{\cos\theta\, d\theta}{\sin^2\theta} \\
&= \int \frac{du}{u^2} \\
&= -\frac{1}{u} + c
\end{aligned}$$

Now we have to reverse our substitutions:

$$\begin{aligned}
\int \frac{dx}{x^2\sqrt{1+x^2}} &= -\frac{1}{\sin\theta} + c \\
&= -\csc\theta + c
\end{aligned}$$

It's not clear how to undo the substitution $x = \tan\theta$. Luckily there is a general method for undoing substitutions like this, which is to go back to thinking of trig functions as ratios of side lengths of a right triangle.

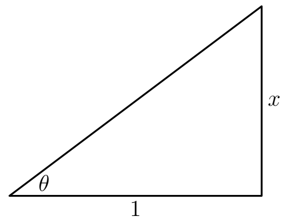


Figure 1: Undoing trig substitution.

We know $x = \tan\theta$ and we know that $\tan\theta$ equals the length of the leg opposite θ divided by the length of the leg adjacent to θ. Figure 1 shows a right triangle with an angle θ, an opposite leg of length x, and an adjacent leg of length 1.

The Pythagorean theorem tells us that the hypotenuse must have length $\sqrt{1+x^2}$. Now we can deduce that:

$$\csc\theta = \frac{\text{hyp}}{\text{opp}} = \frac{\sqrt{1+x^2}}{x}.$$

Hence,

$$\int \frac{dx}{x^2\sqrt{1+x^2}} \quad = \quad -\frac{\sqrt{1+x^2}}{x} + c.$$

In the process of computing this integral we saw the following: trig substitution, rewriting trig functions in terms of sine and cosine, direct substitution, and undoing trig substitution.

What actually happened when we undid that trig substitution was that we computed $\csc(\arctan(x))$. In other words, we composed a trig function with the inverse of another trig function.

Undoing Trig Substitution

Professor Miller plays a game in which students give him a trig function and an inverse trig function, and then he tries to compute their composition. As we've seen, this is sometimes the final step in integration by trig substitution.

$$\_\_(\text{arc}\_\_x) = ?$$

Example: $\tan(\text{arccsc}\, x) = ?$

Question: Isn't $\tan(\text{arccsc}\, x)$ acceptable as a final answer?

Answer: What does "acceptable" mean? The expression $-\csc(\arctan x)$ was a *correct* final answer, but $\dfrac{\sqrt{1+x^2}}{x}$ is a nicer, more insightful, and probably more useful answer.

To simplify $\tan(\text{arccsc}\, x)$ we draw a triangle illustrating an angle whose cosecant is x; see Figure 1. We know that

$$x = \csc\theta = \frac{1}{\sin\theta} = \frac{\text{hyp}}{\text{opp}}$$

so we choose convenient values x and 1 to be the lengths of the hypotenuse and opposite side.

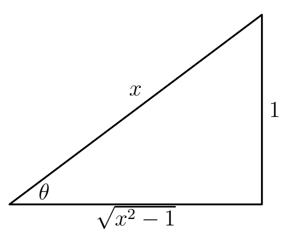


Figure 1: $\theta = \text{arccsc}\, x$ so $x = \csc\theta$.

Once we've drawn our triangle we can compute that the length of the adjacent side must be $\sqrt{x^2-1}$, and so

$$\tan\theta = \frac{\text{opp}}{\text{adj}} = \frac{1}{\sqrt{x^2-1}}.$$

Since $x = \csc\theta$, we have:

$$\tan(\text{arccsc}\, x) = \tan\theta = \frac{1}{\sqrt{x^2-1}}.$$

Whenever you have to undo a trig substitution, this technique is likely to be useful.

Summary of Trig Substitution

Here is a table of different trig substitutions and how they can be useful.

If your integrand contains	Make substitution	To get
$\sqrt{a^2 - x^2}$	$x = a\cos\theta$ or $x = a\sin\theta$	$a\sin\theta$ or $a\cos\theta$
$\sqrt{a^2 + x^2}$	$x = a\tan\theta$	$a\sec\theta$
$\sqrt{x^2 - a^2}$	$x = a\sec\theta$	$a\tan\theta$

These are the three basic forms which are integrated using trig substitution. In general, you use trig substitution to replace the square root of a quadratic function by a trigonometric function. Once you've done this, integrate, then use what we've learned about right triangles and undoing trig substitution to get a final answer.

Completing the Square

The last thing we need to discuss on the topic of trig substitution is completing the square. This is necessary when we want to integrate the square root of a quadratic whose form is not as simple as $ax^2 + c$.

Example: $\displaystyle\int \frac{dx}{\sqrt{x^2 + 4x}}$

The integrand contains the square root of a quadratic, but it doesn't match any of the forms in our table of trig substitutions. What we need to do is use substitution to rewrite it in a form that *is* in our table.

We'll start by trying to write the quadratic in the form:

$$(x + a)^2 + c.$$

$$\begin{aligned} x^2 + 4x &= (x + a)^2 + c \\ &= x^2 + 2xa + a^2 + c \end{aligned}$$

Since the coefficient on x on the left must equal that on the right, it must be true that $2xa = 4x$; i.e. $a = 2$. Similarly, we get $a^2 + c = 0$ or $c = -4$. So:

$$x^2 + 4 = (x + 2)^2 - 4.$$

This process of eliminating the "middle term" using the square of a linear function is called *completing the square*. When we plug this into our integral, it becomes:

$$\int \frac{dx}{\sqrt{x^2 + 4x}} = \int \frac{dx}{\sqrt{(x + 2)^2 - 4}}$$

This is closer to the forms listed in our trig substitution table, but a direct substitution of $u = x + 2$ brings us even closer:

$$\int \frac{dx}{\sqrt{x^2 + 4x}} = \int \frac{du}{\sqrt{u^2 - 4}}$$

We now consult our table and perform the trig substitution $u = 2\sec\theta$, $du = 2\sec\theta\tan\theta\, d\theta$.

$$\begin{aligned} \int \frac{dx}{\sqrt{x^2 + 4x}} &= \int \frac{2\sec\theta\tan\theta\, d\theta}{\sqrt{4\sec^2\theta - 4}} \\ &= \int \frac{2\sec\theta\tan\theta\, d\theta}{2\sqrt{\tan^2\theta}} \\ &= \int \frac{\sec\theta\tan\theta\, d\theta}{\tan\theta} \\ &= \int \sec\theta\, d\theta \\ &= \ln(\sec\theta + \tan\theta) + c \end{aligned}$$

We've completed the integration, but we still need to reverse our two substitutions. We readily see that $\sec\theta = \frac{u}{2}$, and in the process of computing the integral we calculated that $\sqrt{u^2-4} = 2\tan\theta$, so $\tan\theta = \dfrac{\sqrt{u^2-4}}{2}$. Hence we can avoid drawing a triangle and say:

$$\int \frac{dx}{\sqrt{x^2+4x}} = \ln\left(\frac{u}{2} + \frac{\sqrt{u^2-4}}{2}\right) + c$$

Finally, we replace u by $x+2$ to reverse the first substitution:

$$\begin{aligned}\int \frac{dx}{\sqrt{x^2+4x}} &= \ln\left(\frac{x+2}{2} + \frac{\sqrt{(x+2)^2-4}}{2}\right) + c \\ &= \ln\left(\frac{x+2}{2} + \frac{\sqrt{x^2+4x}}{2}\right) + c\end{aligned}$$

Partial Fractions

Today we'll learn how to integrate functions of the form:

$$\frac{P(x)}{Q(x)}$$

where $P(x)$ and $Q(x)$ are polynomials. Functions of this type are called *rational functions*. The technique for integrating functions of this type is called the method of *partial fractions*.

The method of partial fractions works by algebraically splitting $P(x)/Q(x)$ into pieces that are easier to integrate.

Example:

$$\int \left(\frac{1}{x-1} + \frac{3}{x+2} \right) dx = \ln|x-1| + 3\ln|x+2| + c$$

That was an easy integral. Next we'll see how the same integral might become difficult; if we add the two fractions together we get the same problem in a more challenging form:

$$\begin{aligned}
\frac{1}{x-1} + \frac{3}{x+2} &= \frac{1}{x-1} \cdot \frac{x+2}{x+2} + \frac{3}{x+2} \cdot \frac{x-1}{x-1} \\
&= \frac{(x+2) + 3(x-1)}{(x-1)(x+2)} \\
&= \frac{4x-1}{x^2+x-2} \\
\int \frac{4x-1}{x^2+x-2}\, dx &= ?
\end{aligned}$$

This integral is algebraically the same as the one we just computed; that was easy, this one looks harder. The method of partial fractions involves algebraically manipulating integrals like the harder one to make them easier.

Introduction to the Cover-up Method

To integrate a rational function using the partial fractions method we must algebraically break it into parts. We're going to helped in this by a shortcut called the *cover-up method.*

We've disguised the easy integral:

$$\int \left(\frac{1}{x-1} + \frac{3}{x+2} \right) dx$$

as a harder one:

$$\int \frac{4x-1}{x^2+x-2}\, dx.$$

We'll use the method of partial fractions to unwind this disguise.

1. Our first step is to write down the integrand. Then we begin to undo the damage that we did by factoring the denominator (this can be a rather difficult step).

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

2. Next we "set up" some unknowns, preparing to break the rational expression into pieces whose denominators are the factors we just found

$$\frac{4x-1}{(x-1)(x+2)} = \frac{A}{x-1} + \frac{B}{x+2}$$

3. The third step is to solve for the numerators; in this example they are A and B. Once we've completed this we've unwound the disguise.

The cover-up method makes step 3 more efficient and less clumsy.

We solve for A by multiplying by $x-1$.

$$\frac{4x-1}{x+2} = A + \frac{B}{x+2}(x-1)$$

Notice that we didn't try to clear the denominators completely; we just cleared one factor, which was all we needed to do to get A by itself.

If we plug in $x = 1$ we get:

$$\begin{aligned} \frac{4-1}{1+2} &= A + 0 \\ \frac{3}{3} &= A \\ A &= 1 \end{aligned}$$

This is not a surprise — we saw at the start of the lecture that:

$$\frac{4x-1}{x^2+x-2} = \frac{1}{x-1} + \frac{3}{x+2}.$$

Question: Why did you choose $x = 1$?

Answer: Because it works really fast — that's part of the cover-up method. Notice that if we'd set $x = 1$ in the original equation we'd have gotten a denominator of zero; that wouldn't have helped us at all.

What we did was multiply both sides by $x - 1$ and then immediately set $x = 1$; that's like mutiplying both sides by zero! It turns out to be OK though because the equation is true *except* when $x = 1$, so what we're really getting is the limit as x approaches 1.

Setting $x = 1$ helps us by canceling out the term with the variable B in it.

We're going to learn a quicker way to do this in a second, but first let's find the value of B. To isolate B we multiply both sides by $x + 2$:

$$\begin{aligned} \frac{4x-1}{(x-1)(x+2)} &= \frac{A}{x-1} + \frac{B}{x+2} \\ \frac{4x-1}{x-1} &= \frac{A}{x-1}(x+2) + B \end{aligned}$$

Then we plug in $x = -2$:

$$\begin{aligned} \frac{4(-2)-1}{(-2)-1} &= 0 + B \\ \frac{-9}{-3} &= 0 + B \\ B &= 3 \end{aligned}$$

We can now replace A and B by the values we've found to conclude that:

$$\frac{4x-1}{x^2+x-2} = \frac{1}{x-1} + \frac{3}{x+2}.$$

What we've seen here is why the cover-up method works. Next we'll see how we can make it work faster.

Question: Can we use this method if the exponents on x aren't whole numbers? Will it work on square roots?

Answer: This only works with polynomials, which have whole numbered powers of x. For example, x^2+x-2 has powers of 2, 1 and 0 in the denominator. It won't work with square roots.

The Cover-up Method

Here's how the cover-up method works in general:

1. Factor the denominator $Q(x)$,
2. Set-up (describe the target sum),
3. Cover-up (solve for unknown coefficients).

If we can go through these steps without writing writing down all the details of each one, the process will be much faster. It turns out that we can save a lot of time on step 3.

$$\overbrace{\frac{4x-1}{(x-1)(x+2)}}^{\text{Step 1:}} \quad = \quad \overbrace{\frac{A}{x-1}+\frac{B}{x+2}}^{\text{Step 2:}}$$

In Step 3 we cover up the $(x-1)$ in the denominator of $\frac{4x-1}{(x-1)(x+2)}$ and focus on the variable A (which is being divided by $x-1$). Plug the thing that makes $x-1=0$ (i.e. $x=1$) into the part of $\frac{4x-1}{(x-1)(x+2)}$ that's not covered up:

$$\begin{aligned} \frac{4x-1}{\cancel{(x-1)}(x+2)} &= \frac{A}{\cancel{x-1}} + \cancel{\frac{B}{x+2}} \\ \frac{4\cdot 1-1}{1+2} &= A \end{aligned}$$

This is equivalent to the algebra we did earlier, but much faster.

We do the same thing for B, plugging in the value $x=-2$ that makes the denominator associated with B equal to zero:

$$\begin{aligned} \frac{4x-1}{(x-1)\cancel{(x+2)}} &= \cancel{\frac{A}{x-1}} + \frac{B}{\cancel{x+2}} \\ \frac{4(-2)-1}{-2-1} &= B \end{aligned}$$

In general, the cover-up method works if $Q(x)$ has distinct linear factors and the degree of polynomial $P(x)$ is also strictly less than the degree of $Q(x)$. When we have several variables in our set-up, the cover-up method is much more convenient than the step-by-step algebraic solution.

Example: Suppose we're given the rational expression:

$$\frac{x^2+3x+8}{(x-1)(x-2)(x+5)}.$$

The denominator is already factored, so we can skip step 1. From step 2 we get:

$$\frac{x^2+3x+8}{(x-1)(x-2)(x+5)} = \frac{A}{x-1} + \frac{B}{x-2} + \frac{C}{x+5}$$

When using this method, there will always be one term in the final sum for each linear factor in the denominator.

Next we would find the value of each variable, A, B, and C, using the cover-up method.

$$\begin{aligned}
\frac{x^2+3x+8}{\cancel{(x-1)}(x-2)(x+5)} &= \frac{A}{\cancel{x-1}} + \cancel{\frac{B}{x-2} + \frac{C}{x+5}} \\
\frac{1^2+3\cdot 1+8}{(1-2)(1+5)} &= A \\
A &= -2
\end{aligned}$$

Repeated Factors

Now that we've seen the basic method of partial fractions we need to address possible complications. The first complication we'll consider is what to do if the factors in the denominator are not distinct — i.e. if some of the factors are repeated. In order for this technique to work, the degree of the numerator must still be less than the degree of the denominator.

Example: $\dfrac{x^2+2}{(x-1)^2(x+2)}$

Again, step 1 has already been done for us. In the set-up, step 2, we need to add a second term for the second factor of $(x-1)$.

$$\frac{x^2+2}{(x-1)^2(x+2)} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{x+2}$$

In general, when you have $(x-a)^n$ in the denominator you get n corresponding terms in your sum; one for each of the powers $(x-a)^1, (x-a)^2, ..., (x-a)^n$. If the expression were:

$$\frac{x^2+2}{(x-1)^3(x+2)},$$

the setup would need to include another term with $(x-1)^3$ in the denominator.

Question: Why does it have to be squared?

Answer: This is a good question; we'll use an analogy to hint at the answer. The reasoning behind the $(x-1)^2$ is similar to the reasoning behind place value in the decimal expansion of a number. Similarly, we could expand the fraction $\frac{7}{16}$ as:

$$\frac{7}{16} = \frac{0}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4}.$$

Because of the 2^4 in the denominator of $\frac{7}{16}$ we need to use powers of 2 up to 2^4 in the denominator to represent $\frac{7}{16}$ in this way.

When you have repeated factors in the denominator the cover-up method still works, but it doesn't work quite as well. The cover-up method will work for the coefficients B and C but *not* for A.

We start by using the cover-up method to solve for C:

$$\begin{aligned}
\frac{x^2+2}{(x-1)^2\cancel{(x+2)}} &= \cancel{\frac{A}{x-1} + \frac{B}{(x-1)^2}} + \frac{C}{\cancel{x+2}} \\
\frac{(-2)^2+2}{((-2)-1)^2} &= C \\
\frac{6}{9} &= C \\
C &= \frac{2}{3}
\end{aligned}$$

The cover-up method will also work to find B:

$$\begin{aligned}
\frac{x^2+2}{\cancel{(x-1)^2}(x+2)} &= \cancel{\frac{A}{x-1}} + \frac{B}{\cancel{(x-1)^2}} + \cancel{\frac{C}{x+2}} \\
\frac{1^2+2}{1+2} &= B \\
B &= 1
\end{aligned}$$

Let's do that again the slow way to see why it worked:

$$\begin{aligned}
(x-1)^2\frac{x^2+2}{(x-1)^2(x+2)} &= \frac{A}{x-1}(x-1)^2 + \frac{B}{(x-1)^2}x-1)^2 + \frac{C}{x+2}(x-1)^2 \\
\frac{x^2+2}{(x+2)} &= A(x-1) + B + \frac{C}{x+2}(x-1)^2
\end{aligned}$$

When we set $x = 1$, every term with a multiple of $(x-1)$ in it becomes zero and we're left with the value of B.

We can't get everything to cancel so nicely to give us the value of A, which is divided only by a single power of $(x-1)$. If we multiply through by just $(x-1)$ we'll still have an $(x-1)$ in the denominator of the B term which would cause a division by 0. If we multiply through by $(x-1)^2$ the A term cancels completely.

So we have to find another strategy to solve for A. Let's try plugging in Professor Jerison's favorite number, $x = 0$. (Unfortunately, if you use $x = 0$ in solving for other terms in the decomposition you can't use it here. So far we've used $x = 1$ and $x = -2$ so it's ok to use $x = 0$.) Plugging in $x = 0$, $B = 1$ and $C = \frac{2}{3}$, we get:

$$\begin{aligned}
\frac{0^2+2}{(0-1)^2(0+2)} &= \frac{A}{0-1} + \frac{1}{(0-1)^2} + \frac{2/3}{0+2} \\
\frac{2}{2} &= \frac{A}{-1} + \frac{1}{1} + \frac{2/3}{2} \\
1 &= -A + 1 + \frac{1}{3} \\
A &= \frac{1}{3}
\end{aligned}$$

This is a lot of algebra, but if we're careful and thorough we get the right answer and our rational expression becomes easy to integrate.

Question: If $x = 0$ has already been used, what should we do?
Answer: Pick something else, like $x = 1$.

Question: If you had more powers would you have more variables?
Answer: Yes. As the degree of the denominator goes up, the number of variables goes up.

Question: How would you solve it if you had two unknowns?

Answer: When we plug in $x = 0$ (or whatever) we'll get an equation in however many unknowns are left. There are methods of solving systems of equations for those variables which we'll learn more about later.

Quadratic Factors

Our next example is one step more complicated. The degree of the numerator is still less than the degree of the denominator, but the denominator Q will have a quadratic factor:

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

Notice that there's nothing more you can do to factor the denominator.[1]

The set-up for this example looks like:

$$\frac{x^2}{(x-1)(x^2+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1}.$$

The difference is that the numerator corresponding to the factor (x^2+1) is linear, not constant; in general it will be a polynomial with degree one less than the degree of the numerator.

We can still use the cover-up method to solve for A:

$$\begin{aligned}
\frac{x^2}{\cancel{(x-1)}(x^2+1)} &= \frac{A}{\cancel{x-1}} + \cancel{\frac{Bx+C}{x^2+1}} \\
\frac{1^2}{1^2+1} &= A \\
A &= \frac{1}{2}
\end{aligned}$$

It turns out that the best way to solve for B and C is the "slow way"; to clear the denominators completely by multiplying both sides by $(x-1)(x^2+1)$.

$$\begin{aligned}
\frac{x^2}{(x-1)(x^2+1)}((x-1)(x^2+1)) &= \left(\frac{A}{x-1} + \frac{Bx+C}{x^2+1}\right)((x-1)(x^2+1)) \\
x^2 &= A(x^2+1) + (Bx+C)(x-1) \\
x^2 &= Ax^2 + A + Bx^2 + Cx - Bx - C \\
x^2 &= (A+B)x^2 + (C-B)x + (A-C)
\end{aligned}$$

The coefficient of x^2 on the left hand side is 1. On the right hand side, the coefficient of x^2 is $A+B$. We know that $A = \frac{1}{2}$, so we get:

$$\begin{aligned}
1 &= A + B \\
1 &= \frac{1}{2} + B \\
B &= \frac{1}{2}
\end{aligned}$$

[1] If you use complex numbers you can further factor the denominator and again use the cover-up method; in fact, this method was originally intended to be used with complex numbers. This technique was first used by Heaviside in his work with Laplace transforms and inversion of differential equations, but it also turns out to be very useful for integration.

The constant term on the left hand side is 0. On the right hand side the constant term is $A - C = \frac{1}{2} - C$. We conclude that:

$$C = \frac{1}{2}.$$

Question: Why didn't you use the coefficient of x^1?

Answer: It turns out that I didn't need it. From experience, I know that the highest and lowest degree terms of the product will be the easiest terms to work with. I avoid those when possible because they usually involve more arithmetic.

Question: Could you just set $x = 0$?

Answer: Absolutely. That's equivalent to equating the coefficients of the constant term. "Plugging in numbers" is an alternate method of solving these equations, but you still want to avoid plugging in the same number twice.

Question: What if we had $x^3 + 1$ in place of $x^2 + 1$?

Answer: If there's an $x^3 + 1$ in the denominator we haven't fully factored the denominator; we're back to step 1. This method of simplifying rational expressions doesn't work if we don't factor the denominator in the first step. Steps 1 and 2 would become:

$$\begin{aligned}\frac{x^2}{(x-1)(x^3+1)} &= \frac{x^2}{(x-1)(x+1)(x^2-x+1)} \\ &= \frac{A}{x-1} + \frac{B}{x+1} + \frac{Cx+D}{x^2-x+1}\end{aligned}$$

Question: If the degree of the numerator equals the degree of the denominator, none of this works?

Answer: It definitely doesn't work. We're going to have to do something totally different to handle that case.

Remember that the point of all this work was to make it possible to compute an integral:

$$\begin{aligned}\int \frac{x^2}{(x-1)(x^2+1)}\,dx &= \int \left(\frac{A}{x-1} + \frac{Bx+C}{x^2+1}\right) dx \\ &= \int \left(\frac{\frac{1}{2}}{x-1} + \frac{\frac{1}{2}x + \frac{1}{2}}{x^2+1}\right) dx \\ &= \int \frac{\frac{1}{2}}{x-1}\,dx + \int \frac{\frac{1}{2}x}{x^2+1}\,dx + \int \frac{\frac{1}{2}}{x^2+1}\,dx\end{aligned}$$

We've split the integral of the rational expression into three simpler integrals. The first is easy to solve by substituting $u = x - 1$, the second can be solved by "advanced guessing" or by substituting $v = x^2 + 1$. The last one can be solved by trigonometric substitution or by remembering the antiderivative of $\dfrac{1}{x^2+1}$.

$$\int \frac{x^2}{(x-1)(x^2+1)}\,dx = \frac{1}{2}\ln|x-1| + \frac{1}{4}\ln(x^2+1) + \frac{1}{2}\tan^{-1}x + c.$$

Long Division

When you're integrating a rational expression $\dfrac{P(x)}{Q(x}$, what happens if the degree of P is greater than or equal to the degree of Q? We can think of this as an improper fraction, similar to fractions like $\frac{5}{4}$ and $\frac{8}{3}$.

Example:

$$\frac{x^3}{(x-1)(x+2)}$$

The numerator has degree 3 and the denominator has degree 2; our usual method is not going to work here.

The first step to simplifying this turns out to be to reverse our usual step 1; we don't want the denominator factored, we want it expanded, or multiplied out:

$$\frac{x^3}{(x-1)(x+2)} = \frac{x^3}{x^2+x-2}.$$

Our next step is to use long division to convert an improper fraction into a proper fraction:

$$\begin{array}{r@{\,}r}
 & x-1 \\ \cline{2-2}
x^2+x-2\big) & x^3 \qquad\qquad \\
 & -x^3-x^2+2x \quad \\ \cline{2-2}
 & -x^2+2x \quad \\
 & x^2 \;\; +x-2 \\ \cline{2-2}
 & 3x-2
\end{array}$$

You may recall from grade school that $x-1$ is called the *quotient* and $3x-2$ is the *remainder*. We use this result to rewrite our rational expression as follows:

$$\frac{x^3}{(x-1)(x+2)} = \underbrace{x-1}_{\text{easy}} + \underbrace{\frac{3x-2}{x^2+x-2}}_{\text{use cover-up}}$$

We could now find the integral of this rational expression if we wished to.

Partial Fractions – Big Example

We've seen how to do partial fractions in several special cases; now we'll do a big example so that you can see how all these cases fit together.

Remember that partial fractions is a method for breaking up rational expressions into integrable pieces. The good news is, it always works. The bad news is that it can take a lot of time to make it work.

- **Step 0** Long Division:

$$\frac{P(x)}{Q(x)} = \text{quotient} + \frac{R(x)}{Q(x)}.$$

By completing this step you split your rational function into an easy to integrate quotient and a rational function for which the degree of the denominator is greater than the degree of the numerator.

- **Step 1** Factor the Denominator $Q(x)$.

For example, suppose our remainder term looks like:

$$\frac{R(x)}{(x+2)^4(x^2+2x+3)(x^2+4)^3}$$

where the degree of $R(x)$ is less than 12. Polynomials can be extremely difficult to factor; we may need a machine to do this. This can be the hardest step in this method.

If we expand the denominator in the example we get something like:
$x^{12}+10x^{11}+55x^{10}+224x^9+716x^8+1856x^7+4000x^6+7168x^5+10624x^4+12800x^3+12032x^2+8192x+3072$
Factoring this polynomial by hand would be unpleasant.

- **Step 2** Set-up:

$$\frac{R(x)}{(x+2)^4(x^2+2x+3)(x^2+4)^3} = \begin{array}{l} \dfrac{A_1}{(x+2)} + \dfrac{A_2}{(x+2)^2} + \dfrac{A_3}{(x+2)^3} + \\ \dfrac{A_4}{(x+2)^4} + \dfrac{B_0x+C_0}{x^2+2x+3} + \\ \dfrac{B_1x+C_1}{(x^2+4)} + \dfrac{B_2x+C_2}{(x^2+4)^2} + \dfrac{B_3x+C_3}{(x^2+4)^3} \end{array}$$

Note that repeated quadratic factors in the denominator are treated very much the same way as repeated linear factors. There's one term for each power of the repeated factor, and the degree of the numerator is the same in each term.

There are 12 unknowns in this equation; that's not a coincidence. The degree of the denominator is 12. The numerator $R(x)$ can have at most 12 coefficients a_0, a_1, ..., a_{11}; i.e. the number of degrees of freedom of a polynomial of degree 11 is 12.

This is a very complicated system of equations: twelve equations for twelve unknowns. Machines handle this very well, but human beings have a little trouble.

- **Step 3** Cover-up.

  We can use the cover-up method to solve for A_4. That reduces the problem to eleven equations in eleven unknowns.

Question: I see that there are twelve unknowns, but isn't it just one big equation?

Answer: If you multiply both sides by $Q(x)$ it becomes a polynomial equation. On one side you have the known polynomial:

$$R(x) = a_{11}x^{11} + a_{10}x^{10} + ...$$

On the other side of the equation you'll have a polynomial whose coefficients are linear combinations of the unknowns:

$$A_1(x+2)^3(x^2+2x+3)(x^2+4)^3 + A_2(\cdots$$

When you set the coefficients on both sides equal to each other you get 12 equations in 12 unknowns.

Question: Should I write down all this stuff?

Answer: That's a good question! You'll notice that Professor Jerison didn't write it down. It's pages long. You're a human being, not a machine; don't try this at home.

Remember that once we've decomposed $\frac{P(x)}{Q(x)}$ into simpler fractions we still need to integrate it. The quotient and the fractions of the form $\dfrac{A_i}{(x+2)^i}$ are easy to integrate. However, we'll also need to compute something like:

$$\int \frac{x}{(x^2+4)^3}\,dx = -\frac{1}{4}(x^2+4)^{-2} + c$$

using advanced guessing or substitution of $u = x^2 + 4$. To calculate something like:

$$\int \frac{dx}{(x^2+4)^3}$$

we'd need to use the trig substitution $x = 2\tan u$, $dx = 2\sec^2 u\,du$

$$\begin{aligned}
\int \frac{dx}{(x^2+4)^3} &= \int \frac{2\sec^2 u\,du}{(4\sec^2 u)^3} \\
&= \frac{2}{64}\int \cos^4 u\,du \\
&= \frac{1}{32}\int \left(\frac{1+\cos(2u)}{2}\right)^2 du \\
&\vdots
\end{aligned}$$

When calculating:

$$\int \frac{dx}{x^2 + 2x + 3} = \int \frac{dx}{(x+1)^2 + 2},$$

we'll have to complete the square and then use a trig substitution to get something like:

$$\frac{1}{\sqrt{2}} \tan^{-1}\left(\frac{x+2}{\sqrt{2}}\right) + c.$$

In addition, the integral of:

$$\int \frac{x}{x^2 + 2x + 3}\, dx$$

is an expression involving $\ln(x^2 + 2x + 3)$. In theory, we know how to do each of these twelve integrals. In practice it will take a long time.

There's no easier method. The method we use to compute these integrals is going to be at least as complicated as the results, and we've seen that the results can get very complicated. But this method always works, and there are computer programs that can do these calculations for us.

Introduction to Integration by Parts

Unlike the previous method, we already know everything we need to to understand integration by parts. Integration by parts is like the reverse of the product formula:

$$(uv)' = u'v + uv'$$

combined with the fundamental theorem of calculus.

To derive the formula for integration by parts we just rearrange and integrate the product formula:

$$\begin{aligned}
(uv)' &= u'v + uv' \\
uv' &= (uv)' - u'v \\
\int uv'\,dx &= \int (uv)'\,dx - \int u'v\,dx \\
\int uv'\,dx &= uv - \int u'v\,dx
\end{aligned}$$

The integration by parts formula is:

$$\int uv'\,dx = uv - \int u'v\,dx.$$

For definite integrals, it becomes:

$$\int_a^b uv'\,dx = uv\Big|_a^b - \int_a^b u'v\,dx.$$

Example: $\int \ln x\, dx$

This looks intractable, but if we fit it into the form $\int uv'\, dx$, integration by parts makes the calculation relatively easy.

Here's the idea: if we let $u = \ln x$ then when we apply the formula for integration by parts we'll get an integral involving $u' = \frac{1}{x}$. The key element is that the derivative of $u = \ln x$ is easier to integrate than what we started with.

In order to fit the form $\int uv'\, dx$ we need a function v. If we choose $v = x$ then $v' = 1$ and:

$$\int \ln x\, dx = \int uv'\, dx.$$

The formula for integration by parts is:

$$\int uv'\, dx = uv - \int u'v\, dx.$$

So by plugging in $u = \ln x$ and $v = x$ we get:

$$\begin{aligned} \int \underbrace{\ln x}_{uv'}\, dx &= \underbrace{\ln x \cdot x}_{uv} - \int \underbrace{\frac{1}{x}}_{u'} \underbrace{x}_{v}\, dx \\ &= x \ln x - x + c \end{aligned}$$

Example: $\int (\ln x)^2\,dx$

To finish learning the method of integration by parts we just need a lot of practice. To this end, we'll do two slightly more complicated examples.

To integrate:

$$\int (\ln x)^2\,dx,$$

assign:

$$\begin{aligned} u &= (\ln x)^2 \qquad & u' &= 2(\ln x)\frac{1}{x} \\ v &= x \qquad & v' &= 1. \end{aligned}$$

When we differentiate u we get something simpler, which is a good start. Plugging u and v in to the formula for integration by parts we get:

$$\begin{aligned} \int \underbrace{(\ln x)^2}_{uv'}\,dx &= \underbrace{(\ln x)^2 \cdot x}_{uv} - \int \underbrace{2\ln x\frac{1}{x}}_{u'}\underbrace{x}_{v}\,dx \\ &= x(\ln x)^2 - 2\int \ln x\,dx. \end{aligned}$$

We haven't solved the problem, but we're back to the previous case; we recently computed that $\int \ln x\,dx = x\ln x - x + c$. So we have:

$$\int (\ln x)^2\,dx = x(\ln x)^2 - 2\underbrace{(x\ln x - x)}_{\int \ln x\,dx} + c.$$

As we'll see in the next example, this is typical. Integration by parts frequently involves replacing a "hard" integral by an easier one.

A Reduction Formula

When using a *reduction formula* to solve an integration problem, we apply some rule to rewrite the integral in terms of another integral which is a little bit simpler. We may have to rewrite that integral in terms of another integral, and so on for n steps, but we eventually reach an answer.

For example, to compute:

$$\int (\ln x)^n \, dx$$

we repeat the integration by parts from the previous example $n-1$ times, until we're just calculating $\int (\ln x) \, dx$.

For our first step we use:

$$\begin{aligned} u = (\ln x)^n \qquad & u' = n(\ln x)^{n-1}\frac{1}{x} \\ v = x \qquad & v' = 1. \end{aligned}$$

Then:

$$\begin{aligned} \int (\ln x)^n \, dx &= x(\ln x)^n - n\int (\ln x)^{n-1}\frac{1}{x} x \, dx \\ &= x(\ln x)^n - n\int (\ln x)^{n-1} \, dx \end{aligned}$$

So, if:

$$F_n(x) = \int (\ln x)^n \, dx$$

then we've just shown that:

$$F_n(x) = x(\ln x)^n - nF_{n-1}(x).$$

This is an example of a reduction formula; by applying the formula repeatedly we can write down what $F_n(x)$ is in terms of $F_1(x) = \int \ln x dx$ or $F_0(x) = \int 1 \, dx$.

We illustrate the use of a reduction formula by applying this one to the preceding two examples. We start by computing $F_0(x)$ and $F_1(x)$:

$$\begin{aligned} F_0(x) &= \int (\ln x)^0 \, dx = x + c \\ F_1(x) &= x(\ln x)^1 - 1F_0(x) \quad \text{(use reduction formula)} \\ &= x\ln x - x + c \quad \text{(Example 1)} \\ F_2(x) &= x(\ln x)^2 - 2F_1(x) \quad \text{(use reduction formula)} \\ &= x(\ln x)^2 - 2(x\ln x - x) + c \\ &= x(\ln x)^2 - 2x\ln x + 2x + c \quad \text{(Example 2.)} \end{aligned}$$

This is how reduction formulas work in general.

Another Reduction Formula: $\int x^n e^x \, dx$

To compute $\int x^n e^x \, dx$ we derive another reduction formula. We could replace e^x by $\cos x$ or $\sin x$ in this integral and the process would be very similar.

Again we'll use integration by parts to find a reduction formula. Here we choose

$$u = x^n$$

because

$$u' = nx^{n-1}$$

is a simpler (lower degree) function. If $u = x^n$ then we'll have to have

$$v' = e^x, \quad v = e^x.$$

(Note that the antiderivative of v is no more complicated than v' was — another indication that we've chosen correctly.)

On the other hand, if we used $u = e^x$, then $u' = e^x$ would not be any simpler.

Performing the integration by parts we get:

$$\int \underbrace{x^n e^x}_{uv'} \, dx = \underbrace{x^n e^x}_{uv} - \int \underbrace{x^{n-1} e^x}_{u'v} \, dx.$$

If:

$$G_n(x) = \int x^n e^x \, dx$$

then we get the reduction formula:

$$G_n(x) = x^n e^x - nG_{n-1}(x).$$

Let's illustrate this by computing a few integrals. First we directly compute:

$$G_0(x) = \int x^0 e^x \, dx = e^x + c.$$

Now we can use the reduction formula to conclude that:

$$\begin{aligned} G_1(x) &= xe^x - G_0(x) \\ &= xe^x - e^x + c. \end{aligned}$$

So $\int xe^x \, dx = xe^x - e^x + c.$

Question: How do you know when this method will work?

Answer: Good question! The answer is "only through experience and practice". To use this method on an integrand, we need one factor u of the integrand to get simpler when we differentiate and the other factor v not to get more complicated when we integrate.

We've seen how to use integration by parts to derive reduction formulas. We could also find these formulas by advanced guessing — guess what the formula should be and then check it. Either method is valid.

Volume of a Wine Glass: Horizontal Slices

Now we know all of the techniques of integration anyone knows. We'll celebrate by using our new techniques to answer an interesting question.

Find the volume of an exponential wine glass whose bowl is formed by rotating the portion of the graph of $y = e^x$ that joins $(0, 1)$ and $(1, e)$ about the y-axis.

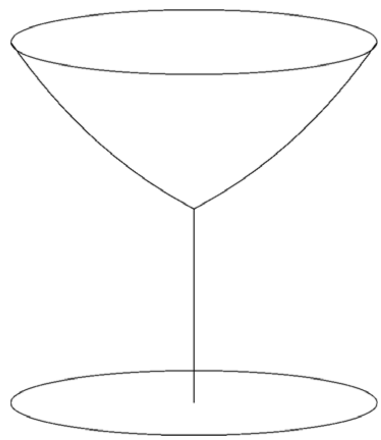


Figure 1: A wine glass formed by rotating the graph of $y = e^x$ about the y-axis.

There are two methods of solving this problem: horizontal and vertical slices. If we compute the volume using horizontal slices we'll be adding up the volumes of disks with height dy and radius $x = \ln y$. (See Figure 2.)

The volume will therefore be:

$$\begin{aligned}
\int_1^e \pi x^2\, dy &= \int_1^e \pi(\ln y)^2\, dy \\
&= \pi F_2(y)\big|_1^e \quad \text{(see previous example)} \\
&= \pi\left[y(\ln y)^2 - 2(y\ln y - y)\right]_1^e \\
&= \pi\left[(e(\ln e)^2 - 2(e\ln e - e)) - (1(\ln 1)^2 - 2(1\ln 1 - 1))\right] \\
&= \pi\left[(e(1)^2 - 2(e\cdot 1 - e)) - (1(0)^2 - 2(1\cdot 0 - 1))\right] \\
&= \pi\left[(e - 2(0)) - (-2(-1))\right] \\
&= \pi(e - 2).
\end{aligned}$$

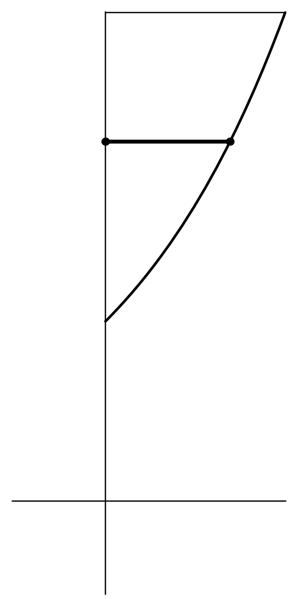


Figure 2: Rotating this horizontal slice about the y-axis forms a disk.

Volume of a Wine Glass: Vertical Slices

If we use vertical slices to compute the volume of our exponential wine glass, we'll be adding up volumes of shells with height $e - y$, radius x and thickness dx.

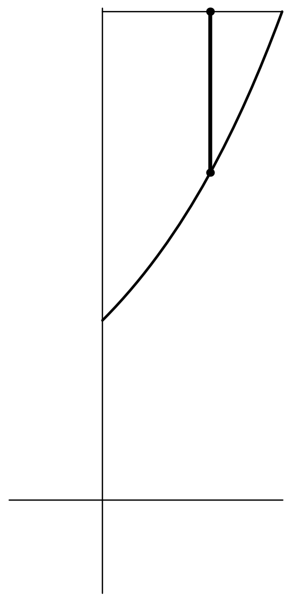


Figure 1: Rotating a slice of thickness dx about the y-axis produces a shell.

$$\begin{aligned}
\text{Volume} &= \int_0^1 (e-y)2\pi x\, dx \\
&= \int_0^1 (e-e^x)2\pi x\, dx \\
&= \int_0^1 2\pi e x\, dx - \int_0^1 2\pi x e^x\, dx \\
&= \underbrace{\frac{2\pi e}{2}}_{\text{area of a triangle}} - 2\pi G_1(x)\Big|_0^1 \\
&= \pi e - 2\pi \left[x e^x - e^x\right]_0^1 \\
&= \pi e - 2\pi \left[(e-e) - (0-1)\right] \\
&= \pi e - 2\pi \\
&= \pi(e-2).
\end{aligned}$$

Introduction to Arc Length

Now that we're done with techniques of integration, we'll return to doing some geometry; this will lead to some of the tools you'll need in multivariable calculus. Our first topic is *arc length*, which is calculated using another cumulative sum which will have an associated story and picture.

Suppose you have a roadway with mileage markers $s_0, s_1, s_2, ..., s_n$ along the road. The distance traveled along the road — the *arc length* — is described by this parameter s. If we look at the road as a graph, we can let a be the x coordinate of the first point s_0 on the curve or road and b be the x coordinate of the end point s_n of the curve, and x_i as the x-coordinate of s_i. This is reminiscent of what we did with Riemann sums.

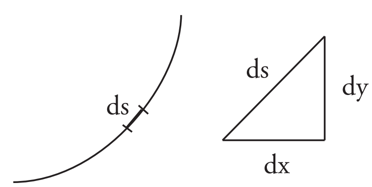


Figure 1: Straight line approximation of arc length.

We'll approximate the length s of the curve by summing the straight line distances between the points s_i. As n increases and the distance between the s_i decreases, the straight line distance from s_i to s_{i-1} will get closer and closer to the distance Δs along the curve. We can use the Pythagorean theorem to see that that distance equals $\sqrt{(\Delta x)^2 + (\Delta y)^2}$. In other words:

$$(\Delta s)^2 \approx \overbrace{(\Delta x)^2 + (\Delta y)^2}^{(\text{hypotenuse})^2}.$$

We apply the tools of calculus to this estimate; in the infinitesimal this is exactly correct:

$$(ds)^2 = (dx)^2 + (dy)^2.$$

In the future we'll omit the parentheses and write this as $ds^2 = dx^2 + dy^2$. These are squares of differentials; try not to mistake them for differentials of squares.

The next thing we do is take the square root:

$$ds = \sqrt{dx^2 + dy^2}.$$

This is the formula that Professor Jerison has memorized, but you can rewrite it in several useful ways. For instance, you can factor out the dx to get:

$$\frac{ds}{dx} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}.$$

This is the form we'll be using today; when we add up all the infinitesimal values of ds we'll find that:

$$\begin{aligned}
\text{Arc Length} &= \text{distance along the curve from } s_0 \text{ to } s_n \\
&= \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \\
&= \int ds \\
&= \int_a^b \sqrt{1 + f'(x)^2}\, dx \quad (y = f(x))
\end{aligned}$$

Question: Is $f'(x)^2$ equal to $f''(x)$?

Answer: No. Suppose $f(x) = x^2$. Then $f'(x) = 2x$, $f'(x)^2 = 4x^2$ and $f''(x) = 2$.

Question: What are the limits of integration on $\int ds$, above?

Answer: If you're integrating with respect to s you'll start at s_0 and end at s_n. If you're integrating with respect to a different variable you'll have different limits of integration, as happens when we change variables. The values s_0 and s_n are mileage markers along the road; they're not the same as a and b. When we measure arc length, remember that we're measuring distance along a curved path.

Example: $y = mx$

This is a basic example that should help you get some perspective on this method.

$$\begin{aligned} y &= mx \\ y' &= m \\ ds &= \sqrt{1+(y')^2}\,dx \\ &= \sqrt{1+m^2}\,dx \end{aligned}$$

The length of the curve $y = mx$ on the interval $0 \leq x \leq 10$ is:

$$\int_0^{10} \sqrt{1+m^2}\,dx = 10\sqrt{1+m^2}$$

We've drawn a picture to confirm this; see Figure 1. We see that the

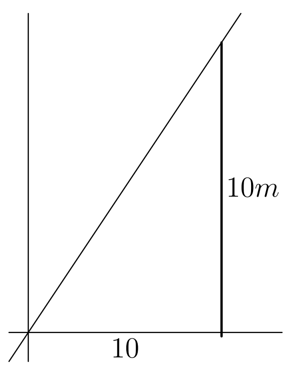


Figure 1: Arc Length of $y = mx$ over $0 \leq x \leq 10$.

arc whose length we're computing is the hypotenuse of a triangle, and the Pythagorean theorem tells us that its length is:

$$\sqrt{10^2 + (10m)^2} = 10\sqrt{1+m^2}.$$

You may be disdainful of how obvious this is. But if we can figure out these formulas for linear functions, we can use calculus to subdivide other functions into infinitesimal linear parts and then solve the problem for those functions. This is the main point of these integrals.

Example: Circular Arc

$$y = \sqrt{1 - x^2}$$

describes the graph of a semicircle. We'll find the arc length of the piece of this semicircle above the interval $0 \le x \le a$. (See Figure 1.)

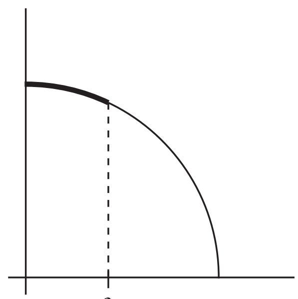


Figure 1: Arc length of $y = \sqrt{1 - x^2}$ over $0 \le x \le a$.

We'll use the variable α to denote the arc length along the circle. We could calculate the exact value of α using trigonometry, but we'll first find it using calculus. We start by finding y'.

$$\begin{aligned} y' &= \frac{-x}{\sqrt{1 - x^2}} \\ ds &= \sqrt{1 + (y')^2}\, dx \\ &= \sqrt{1 + \left(\frac{-x}{\sqrt{1 - x^2}}\right)^2}\, dx \end{aligned}$$

Yuck. Let's simplify $1 + \left(\frac{-x}{\sqrt{1 - x^2}}\right)^2$ in a separate calculation:

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

Plugging this in and using our formula for arc length, we get:

$$\alpha = \int_0^a \frac{dx}{\sqrt{1 - x^2}}$$

$$
\begin{aligned}
&= \sin^{-1} x\Big|_0^a \\
\alpha &= \sin^{-1} a. \\
(\text{So } \sin\alpha &= a.)
\end{aligned}
$$

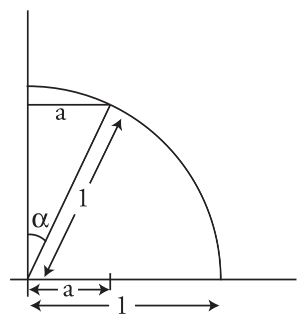


Figure 2: The arc length equals α.

This is a little deeper than it looks; we went a distance α along the arc of a circle that has radius 1 and ended up at a point whose x-coordinate was a.

Previously, if we had an angle whose measure was α radians, we'd say:

$$\sin\alpha = a.$$

You may have been told that radians measured the arc length along the curve of the circle, but this is the first time you've been able to derive it.

Remember that our first definition of the exponential function e^x involved the slope of its graph, but later we were able to define the natural log function as an integral. The same sort of thing is happening here; if you want to know what radians are you have to calculate this arc length. This gives you a new definition of the arcsine function, which gives you a new definition of the sine function, which leads to an improved definition and understanding of trig functions.

Example: Length of a Parabola

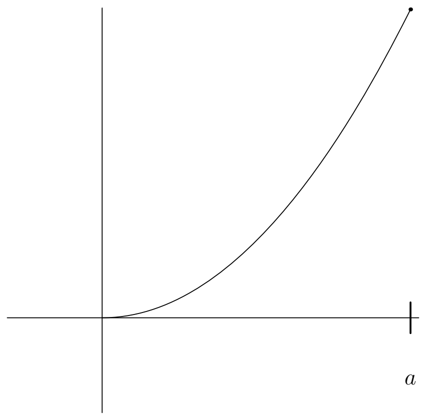


Figure 1: Arc length of $y = x^2$ over $0 \le x \le a$.

To find the arc length of a parabola we start with:

$$\begin{aligned} y &= x^2 \\ y' &= 2x \\ ds &= \sqrt{1+(2x)^2}\,dx \\ &= \sqrt{1+4x^2}\,dx. \end{aligned}$$

So the arc length of the parabola over the interval $0 \le x \le a$ is:

$$\int_0^a \sqrt{1+4x^2}\,dx.$$

This is the answer to the question, but it would be more useful to us if we could write it in a simpler form. That's why we studied techniques of integration. To evaluate this integral we use the following trig substitution:

$$\begin{aligned} x &= \frac{1}{2}\tan u \\ dx &= \frac{1}{2}\sec^2 u \end{aligned}$$

When we do, we find that:

$$\int_0^a \sqrt{1+4x^2}\,dx = \left[\frac{1}{4}\ln(2x+\sqrt{1+4x^2}) + \frac{1}{2}x\sqrt{1+4x^2}\right]_0^a$$

(you may have seen parts of this calculation in a recitation video).

Introduction to Surface Area

We're going to move to three dimensions now to talk about surface area; we'll be doing a lot with surface area in multivariable calculus. If this starts to look too complicated, keep in mind that all we're doing is integrating infinitesimal pieces of simple, linear functions.

The only surface areas we'll compute in this class are surfaces of rotation. We'll start by rotating the parabola from our last example about the x-axis to get the trumpet shape shown in Figure 2. (Remember that we're only interested in the surface — the metal part of the trumpet — and not the interior.)

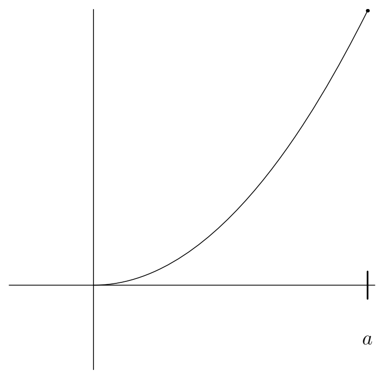


Figure 1: The parabola $y = x^2$.

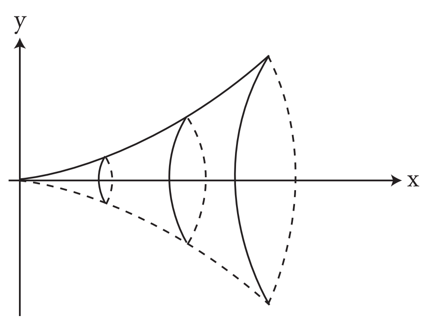


Figure 2: The parabola $y = x^2$ rotated about the x-axis.

We figure out the formula for surface area of a surface of rotation in much the same way we figured out the formula for volumes of revolution. Think of a small segment of arc length with length ds. If that segment is parallel to the x-axis, when you rotate it around the axis it sweeps out a shell shape. If the segment is tilted at an angle, then the surface swept out will have more area than a shell, proportional to the amount of tilt. The surface area swept out is proportional to the length ds of the segment.

In our example, the total surface area swept out by a small segment of arc will be:

$$dA = \underbrace{(2\pi y)}_{\text{circumference}} (ds).$$

You may also see S used for surface area (and s used for arc length):

$$dS = (2\pi y)(ds).$$

The surface area of our trumpet shape will then be:

$$\begin{aligned} \text{Surface area} &= \int_0^a \underbrace{2\pi x^2}_{2\pi y} \underbrace{\sqrt{1+4x^2}\,dx}_{ds \text{ from before}} \\ &\vdots \qquad\qquad (\text{substitute } x = \tfrac{1}{2}\tan u) \end{aligned}$$

The calculation for this integral is long; if we wanted to we could use a computer program to get an answer. Our goal is to be able to see that we could get an exact solution to this integral if we had to; i.e. that we could rewrite it as a product of trigonometric functions.

Surface Area of a Sphere

In this example we will complete the calculation of the area of a surface of rotation. If we're going to go to the effort to complete the integral, the answer should be a nice one; one we can remember. It turns out that calculating the surface area of a sphere gives us just such an answer.

We'll think of our sphere as a surface of revolution formed by revolving a half circle of radius a about the x-axis. We'll be integrating with respect to x, and we'll let the bounds on our integral be x_1 and x_2 with $-a \leq x_1 \leq x_2 \leq a$ as sketched in Figure 1.

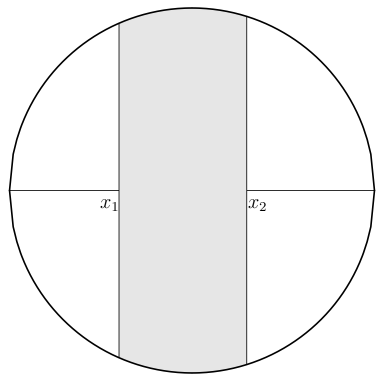


Figure 1: Part of the surface of a sphere.

Remember that in an earlier example we computed the length of an infinitesimal segment of a circular arc of radius 1:

$$ds = \sqrt{\frac{1}{1-x^2}}\,dx$$

In this example we let the radius equal a so that we can see how the surface area depends on the radius. Hence:

$$\begin{aligned} y &= \sqrt{a^2 - x^2} \\ y' &= \frac{-x}{\sqrt{a^2 - x^2}} \\ ds &= \sqrt{1 + \frac{x^2}{a^2 - x^2}}\,dx \\ &= \sqrt{\frac{a^2 - x^2 + x^2}{a^2 - x^2}}\,dx \\ &= \sqrt{\frac{a^2}{a^2 - x^2}}\,dx. \end{aligned}$$

The formula for the surface area indicated in Figure 1 is:

$$\text{Area} = \int_{x_1}^{x_2} 2\pi y\,ds$$

$$
\begin{aligned}
&= \int_{x_1}^{x_2} 2\pi \overbrace{\sqrt{a^2 - x^2}}^{y} \overbrace{\sqrt{\frac{a^2}{a^2 - x^2}}\,dx}^{ds} \\
&= \int_{x_1}^{x_2} 2\pi \sqrt{a^2 - x^2} \frac{a}{\sqrt{a^2 - x^2}}\,dx \\
&= \int_{x_1}^{x_2} 2\pi a\,dx \\
&= 2\pi a(x_2 - x_1).
\end{aligned}
$$

Special Cases

When possible, we should test our results by plugging in values to see if our answer is reasonable. Here, if we set $x_1 = 0$ and $x_2 = a$ we should get the surface area of a hemisphere of radius a:

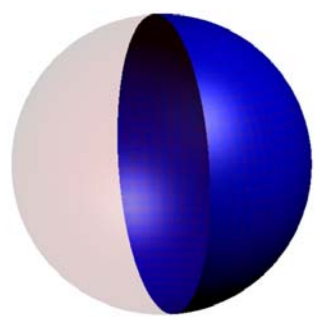


Figure 2: Right hemisphere.

$$
\begin{aligned}
2\pi a(x_2 - x_1) &= 2\pi a(a - 0) \\
&= 2\pi a^2
\end{aligned}
$$

We get the surface area of the whole sphere by letting $x_1 = -a$ and $x_2 = a$:

$$
\begin{aligned}
2\pi a(x_2 - x_1) &= 2\pi a(a - (-a)) \\
&= 4\pi a^2
\end{aligned}
$$

Question: Would it be possible to rotate around the y-axis?

Answer: Yes. If we rotate around the y axis and integrate with respect to x (calculating the surface area of a vertical slice, as we did here) we'd be adding up little strips of area. If we integrate with respect to y and find the surface area between two vertical positions y_1 and y_2 we'd get exactly the same calculation.

Question: Can you compute surface area using shells?

Answer: The short answer is "not quite". We use the word shell to describe something which has a thickness dx. Shells have volume, integrals which involve shells compute volumes, not surface areas.

To compute surface area you need to sum up the areas of small regions of your surface, but those small regions can have any shape whose area you can measure.

Parametric Curve

We're going to continue to work in three dimensional space, moving on to parametric equations; in particular we'll discuss parametric curves. This is another topic that will help us prepare for multivariable calculus; it's the beginning of the transition to multivariable thinking.

We're going to consider curves that are described by x being a function of t and y being a function of t.

$$\begin{aligned} x &= x(t) \\ y &= y(t). \end{aligned}$$

The variable t is called a *parameter*. The easiest way to think of parametric curves is as t equaling time and the position $(x(t), y(t))$ describing a *trajectory* in the plane.

The point $(x(0), y(0))$ describes a position at time $t = 0$. The point $(x(1), y(1))$ describes a later position at time $t = 1$. When we draw the trajectory it's a good idea to draw arrows on the curve that show what direction $(x(t), y(t))$ moves in as t increases.

Our first example will be to figure out what sort of curve is described by:

$$\begin{aligned} x &= a\cos t \\ y &= a\sin t. \end{aligned}$$

To do this we want to figure out what equation describes the curve in rectangular coordinates. Ideally, we quickly realize that:

$$\begin{aligned} x^2 + y^2 &= (x(t))^2 + (y(t))^2 \\ &= a^2\cos^2 t + a^2\sin^2 t \\ x^2 + y^2 &= a^2. \end{aligned}$$

The curve is a circle with radius a.

Another thing to keep track of is which direction we're going on this circle. There's more to this curve than just its shape; there's also where we are at what time, as with the trajectory of a planet. We'll figure this out by plotting a few points:

When $t = 0$, $(x, y) = (a\cos 0, a\sin 0) = (a, 0)$.

When $t = \frac{\pi}{2}$, $(x, y) = (a\cos\frac{\pi}{2}, a\sin\frac{\pi}{2}) = (0, a)$.

We deduce that the trajectory moves counterclockwise about the circle of radius a centered at the origin. (See Figure 1.)

Next time we'll learn to keep track of the arc length and to understand how fast $(x(t), y(t))$ is changing.

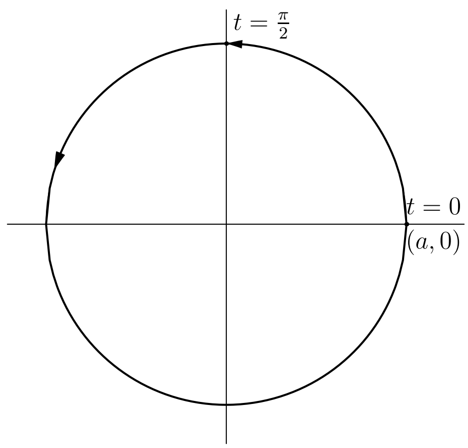


Figure 1: Parametrized circle.

Arc Length of Parametric Curves

We've talked about the following parametric representation for the circle:

$$\begin{aligned} x &= a \cos t \\ y &= a \sin t \end{aligned}$$

We noted that $x^2 + y^2 = a^2$ and that as t increases the point $(x(t), y(t))$ moves around the circle in the counterclockwise direction.

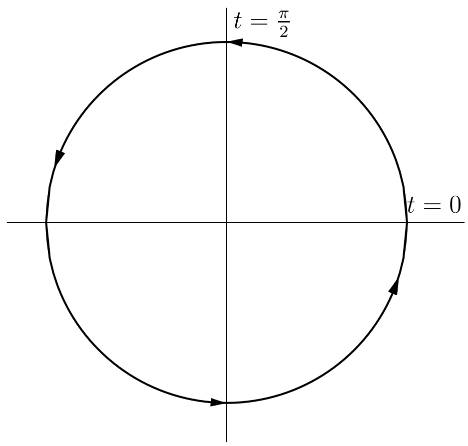


Figure 1: The parametrization $(a \cos t, a \sin t)$ has a counterclockwise trajectory.

We'll now learn how to compute the arc length of the path traced out by this trajectory; the result should match our previous result for the arc length of a circular curve.

Recall our basic relationship:

$$ds^2 = dx^2 + dy^2 \quad \text{or} \quad ds = \sqrt{dx^2 + dy^2}.$$

We incorporate parameter t into this formula as follows:

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

So, to compute the infinitesimal arc length ds we start by computing $\frac{dx}{dt}$ and $\frac{dy}{dt}$:

$$\frac{dx}{dt} = -a \sin t \quad \text{and} \quad \frac{dy}{dt} = a \cos t.$$

Hence,

$$ds = \sqrt{(-a \sin t)^2 + (a \cos t)^2}\, dt$$

$$\begin{aligned} &= \sqrt{a^2(\sin^2 t + \cos^2 t)}\, dt \\ &= \sqrt{a^2 \cdot 1}\, dt \\ ds &= a\, dt \end{aligned}$$

From this we conclude that the speed at which the point moves around the circle is: $\frac{ds}{dt} = a$. Because the speed is constant, we say that the point is moving with *uniform* speed.

Parametrizations such as:

$$\begin{aligned} x &= a \cos kt \\ y &= a \sin kt \end{aligned}$$

are common in math and physics classes. Again this is a parametrization of the circle, but this time the point is moving with uniform speed ak. (We'll assume that both a and k are positive.)

Remarks on Notation

We've been working with notation like ds^2 for a while now; what does this mean, what operations can we legitimately perform with these infinitesimals, and what isn't valid?

The basis for our arc length formula is that:

$$\Delta s^2 \approx \Delta x^2 + \Delta y^2.$$

We'll now see how our formula:

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

for parametric arc length can be more rigorously derived from the same basis.

Because Δt is not quite equal to 0, we can start by dividing both sides of the formula by Δt^2:

$$\begin{aligned} \Delta s^2 &\approx \Delta x^2 + \Delta y^2 \\ \left(\frac{\Delta s}{\Delta t}\right)^2 &\approx \left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2 \end{aligned}$$

Finally, we take the limit as t goes to zero of both sides to conclude that:

$$\left(\frac{ds}{dx}\right)^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2.$$

(This is what derivatives are all about.)

Warning: Never write $\left(\frac{dx}{dt}\right)^2 = (x'(t))^2$ as $\frac{dx^2}{dt^2}$. If you do, it could be incorrectly interpreted to mean $\frac{d^2x}{dt^2} = x''(t)$.

Another unfortunate thing is that we write $\sin^2 x$ to mean $(\sin x)^2$, perhaps because typographers are lazy. There is inconsistency in mathematical notation, and we have to work with the conventions that exist.

Non-Constant Speed Parametrization

Let's look at the following parametrization:

$$\begin{aligned} x &= 2\sin t \\ y &= \cos t. \end{aligned}$$

To solve this sort of problem we're going to need to convert this parametrization in terms of cosine, sine and t into a rectangular equation in x and y. (We'll see a few more examples of this process later today in a different context.)

To see the pattern here we'll use the relationship:

$$\sin^2 t + \cos^2 t = 1.$$

Since $\frac{x}{2} = \sin t$, we get that:

$$\frac{1}{4}x^2 + y^2 = \sin^2 t + \cos^2 t = 1.$$

So the trajectory described by this parametrization is an ellipse.

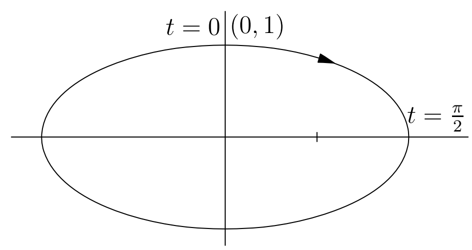


Figure 1: Ellipse described by $x = 2\sin t$, $y = \cos t$.

To sketch the ellipse we'll start by plotting a few points. When $t = 0$ we have:

$$(2\sin 0, \cos 0) = (0, 1),$$

so the ellipse "starts" at the point $(0, 1)$. When $t = \pi/2$ we get the point

$$\left(2\sin\frac{\pi}{2}, \cos\frac{\pi}{2}\right) = (2, 0).$$

We know that the trajectory follows an ellipse, and we can compute that the length of the minor axis is 1 and the major axis is 2, so we get the curve shown in Figure 1.

We also know that the motion is clockwise.

Next, let's examine the speed at which the point traces out arc length. We know:

$$\begin{aligned} \frac{ds}{dt} &= \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \\ &= \sqrt{(2\cos t)^2 + (-\sin t)^2} \end{aligned}$$

And so the total arc length covered by the point as it moves all the way around the ellipse (t varies from 0 to 2π) is:

$$\text{Arc length} = \int_0^{2\pi} \sqrt{4\cos^2 t + \sin^2 t}\, dt.$$

Unfortunately, this is not an elementary integral; we won't be able to find a formula for an antiderivative of $\sqrt{4\cos^2 t + \sin^2 t}$. That means we have to stop here; this is our final answer.

Note that it may be hard to tell whether or not it's possible to find a given antiderivative.

Question: When you draw the ellipse, don't you need to take into account what t is?

Answer: Good question. Our problem is to plot the curve parametrized by:

$$\begin{aligned} x &= 2\sin t \\ y &= \cos t. \end{aligned}$$

Obviously we can't simply graph y as a function of x; both y and x are functions of t. To draw this curve we did two things:

1. Plot points $(x(t), y(t))$ for "easy" values of t. (Here $t = 0$ and $t = \frac{\pi}{2}$.)
2. Find a relationship between x and y similar to one we can graph.

A calculator or computer would draw the curve by repeating the first step many, many times. We're not so patient, so we used the fact that $\left(\frac{1}{2}x\right)^2 + y^2 = 1$ to deduce that the curve must have an elliptical shape. Even if you don't recognize this as the equation of an ellipse, you should be able to guess that its graph will be a deformed circle.

By combining the information we have about locations of points on the curve (we might also want to find points for $t = \pi$ and $t = \frac{3\pi}{2}$) with the information that the curve looks like a deformed circle, we can get a fairly decent sketch of the parametrized curve.

You might be asked to give the rectangular equation for a parametric curve and then plot the curve. In this case, the answer would be $\frac{1}{4}x^2 + y^2 = 1$ followed by a picture of the ellipse.

Question: Do I need to know any specific formulas?

Answer: Any formulas that you know and remember may help you. You're not required to memorize the general equation of an ellipse, but you should recognize the equation of a circle.

Question: Arc length is the integral of $ds = \sqrt{dx^2 + dy^2}$. With ds, dy and dx all in the mix, how were you able to integrate just with respect to x in some examples?

Answer: When working with one dimensional objects like curves in space or in the plane, we're going to integrate with respect to a single variable. We get to choose which variable to use, but some choices are better than others.

For instance, if we're trying to find the arc length of a circle or the ellipse we just looked at, integrating with respect to x might be a problem because there are two different y values for most choices of x. However, if we can solve the problem by looking only at the top half of the circle or ellipse it could be ok to integrate with respect to x.

The uniform parameter (the one for which the point moves with uniform speed) may be the easiest one. In the previous example, the uniform parameter was t.

This returns us to the point that we're no longer tied to the view that y is a function of x. The variables x and y just represent the horizontal and vertical position of a point; they're no longer the input and output of a function. In fact, in this example both x and y are functions of t.

Surface Area of an Ellipsoid

Next we'll find the surface area of the surface formed by revolving our elliptical curve:

$$\begin{aligned} x &= 2\sin t \\ y &= \cos t \end{aligned}$$

about the y-axis.

Remember that our surface area element dA is the area of a thin circular ribbon with width ds. The radius of this circle is $x = 2\sin t$, which is the distance between the ribbon and the y-axis.

$$dA = 2\pi \underbrace{(2\sin t)}_{x} \underbrace{\sqrt{4\cos^2 t + \sin^2 t}dt}_{ds=\text{arc length}}.$$

To find the surface area we need to integrate dA between certain limits; what are they?

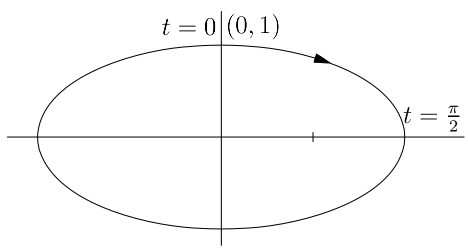


Figure 1: Elliptical path described by $x = 2\sin t$, $y = \cos t$.

By looking at Figure 1 we can see that we need to integrate from 0 to π. Remember that we only need to go from the top to the bottom of the ellipse to trace the right hand side; including the left hand side of the ellipse would double our result and give the wrong answer.

$$A = \int_0^\pi 2\pi(2\sin t)\sqrt{4\cos^2 t + \sin^2 t}dt.$$

Notice that we're integrating from the top of the ellipse to the bottom; if we think in terms of the y-variable we tend to think of going the opposite way.

This integral turns out to be do-able but long. Start by using the substitution $u = \cos t$, $du = -\sin t\, dt$.

$$\begin{aligned} A &= \int_0^\pi 2\pi(2\sin t)\sqrt{4\cos^2 t + \sin^2 t}dt \\ &= \int_{u=1}^{u=-1} -4\pi\sqrt{3u^2+1}du. \end{aligned}$$

Next would be another trigonometric substitution to deal with the square root, and so on.

Introduction to Polar Coordinates

Polar coordinates involve the geometry of circles. Just as Professor Jerison loves the number zero, the rest of MIT loves circles.

Polar coordinates are another way of describing points in the plane. Instead of giving x and y coordinates, we'll describe the location of a point by:

- r = distance to origin
- θ = angle between the ray from the origin to the point and the horizontal axis.

(This is the geometric idea, but is not a perfect match for how polar coordinates are actually used.)

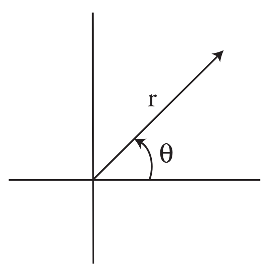


Figure 1: Polar coordinates describe a radius r and angle θ.

If we wish to relate polar coordinates back to rectangular coordinates (i.e. find the x and y coordinates of a point (r, θ)), we use the following formulas:

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

This is the official, unambiguous definition of polar coordinates, from which we get the geometric description above and also the following:

To convert rectangular coordinates to polar coordinates, use:

$$r = \sqrt{x^2 + y^2}, \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right).$$

This is close to being a good formula, and it's useful.

The ambiguity in these formulas comes from the fact that r could be negative $\sqrt{x^2 + y^2}$, and θ could also be $\tan^{-1}\left(\frac{-y}{-x}\right)$. You must refer to your diagram when using these formulas to convert from rectangular to polar coordinates.

Two Coordinate Systems

The coordinate system that we're used to is the rectangular coordinate system. The notation (x, y) describes a location in that plane that is x units to the right of the origin and y units above the origin. As shown in Figure 2, the (green) lines $y = k$ are lines of constant height; the (red) lines $x = c$ are made up of all the points that are exactly c units to the right of the origin.

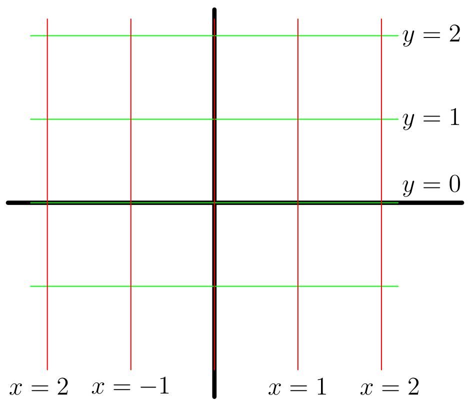


Figure 2: Lines $y = k$ and $x = c$ in a rectangular coordinate system.

In the polar coordinate system, the notation (r, θ) describes a point r units away from the origin at an angle of θ degrees. In Figure 3, each ray $\theta = c$ radiating from the origin is made up of points (r, θ) which all have the same angle θ. The circles $r = k$ about the origin are made up of points which are all the same distance k from the origin.

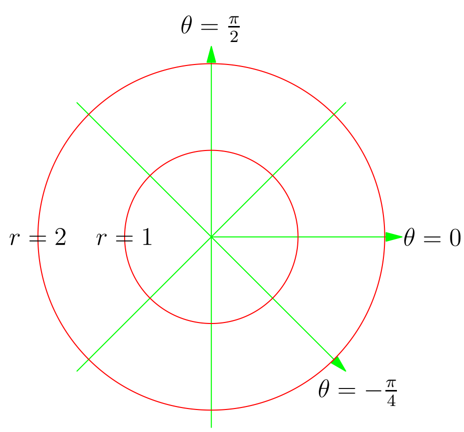


Figure 3: Lines $r = k$ and $\theta = c$ in a polar coordinate system.

The polar coordinate system is just a different way of describing the locations of points in the plane.

Simple Examples in Polar Coordinates

We've just learned about the polar coordinate system, which is very useful in multivariable calculus and in physics. Here are some examples to help you get used to it.

Example: $(x, y) = (1, -1)$

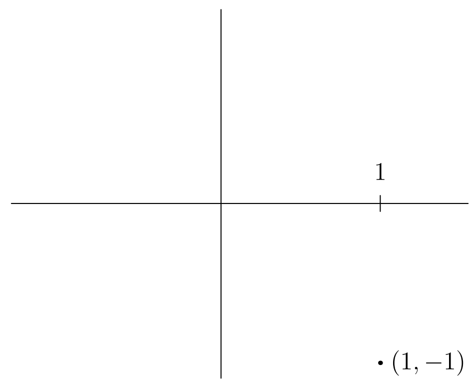


Figure 1: Point at $(1, 1)$ in rectangular coordinates.

How do you describe this point in polar coordinates? There's more than one right answer.

1. $r = \sqrt{2}, \qquad \theta = \frac{7\pi}{4}$
2. $r = \sqrt{2}, \qquad \theta = -\frac{\pi}{4}$
3. $r = -\sqrt{2}, \qquad \theta = \frac{3\pi}{4}$

Some of these answers are easier to make sense of than others, but they are all "legal" correct answers. This ambiguity is something you'll have to adapt to when working with polar coordinates.

Question: Don't the radii have to be positive because they represent a *distance* from the origin?

Answer: When we said r was the distance to the origin in polar coordinates, that was a lie. The only unambiguous description of polar coordinates is:

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

All the others are flawed in some way, but still useful.

It's useful to think of r as a distance, but it's not always accurate to think this way.

Example: $r = a$

In polar coordinates, $r = a$ describes a circle of radius a centered at the origin.

Example: $\theta = c$

The equation $\theta = c$ describes a ray in polar coordinates.

Warning: This implicitly assumes that $0 \leq r < \infty$. If we instead assume $-\infty < r < \infty$ we'd get a line, not a ray.

Typical Conventions in Polar Coordinates

- $0 \leq r < \infty$
- $-\pi < \theta \leq \pi$ **or** $0 \leq \theta < 2\pi$.

These are typical, but not universal; different assumptions and ranges are used in different contexts.

Translating $y = 1$ into Polar Coordinates

We'll take a simple description from rectangular coordinates, $y = 1$, and translate it into polar coordinates. To do this, we plug in the (definitive) formula $y = r \sin\theta$.

$$\begin{aligned} y &= r\sin\theta \\ 1 &= r\sin\theta \\ r &= \frac{1}{\sin\theta} \end{aligned}$$

In rectangular coordinates the line has equation $y = 1$. In polar coordinates its equation is $r = \dfrac{1}{\sin\theta}$.

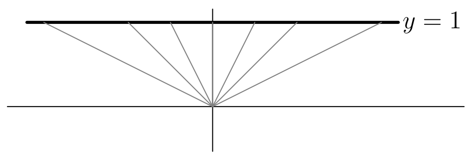


Figure 1: $r = \dfrac{1}{\sin\theta}$

As indicated in Figure 1, for different values of θ points on the horizontal line are different distances r from the origin. That distance r is $\frac{1}{\sin\theta}$.

We need one more piece of information to complete this problem; what is the range of θ? When $\theta = 0$ the denominator of the expression describing r is 0; this corresponds to one end of the line. As θ increases from 0 to π, r decreases to 1 at $\theta = \frac{\pi}{2}$ and then increases to infinity again.

Our final answer is:

$$r = \frac{1}{\sin\theta}, \qquad 0 < \theta < \pi.$$

Question: Is it typical to express r as a function of θ? Does it matter?

Answer: In this course our answers will almost always describe r as a function of θ, but it's not required. We do it this way because we like:

$$r = \frac{1}{\sin\theta}$$

better than:

$$\theta = \sin^{-1}\left(\frac{1}{r}\right).$$

Equation of an Off-Center Circle

This is a standard example that comes up a lot. Circles are easy to describe, unless the origin is on the rim of the circle. We'll calculate the equation in polar coordinates of a circle with center $(a, 0)$ and radius $(2a, 0)$. You should expect to repeat this calculation a few times in this class and then memorize it for multivariable calculus, where you'll need it often.

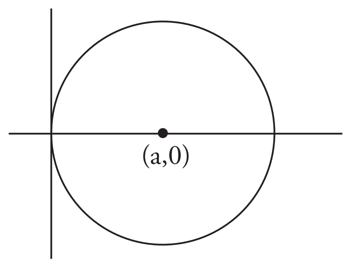


Figure 1: Off center circle through $(0, 0)$.

In rectangular coordinates, the equation of this circle is:

$$(x - a)^2 + y^2 = a^2.$$

We could plug in $x = r\cos\theta$, $y = \sin\theta$ to convert to polar coordinates, but there's a faster way. We start by expanding and simplifying:

$$\begin{aligned}
(x-a)^2 + y^2 &= a^2 \\
x^2 - 2ax + a^2 + y^2 &= a^2 \\
x^2 - 2ax + y^2 &= 0 \\
(x^2 + y^2) - 2ax &= 0 \\
r^2 - 2ar\cos\theta &= 0 \\
r^2 &= 2ar\cos\theta \\
\Longrightarrow r &= 2a\cos\theta \quad (\text{or } r = 0).
\end{aligned}$$

We used the facts that $x^2 + y^2 = r^2$ and $x = r\cos\theta$ to conclude that there were two values of r that satisfy this equation; $r = 2a\cos\theta$ and $r = 0$. These are the equations describing r in terms of θ that describe this circle in polar coordinates.

In order to use the equation $r = 2a\cos\theta$, we need to figure out the appropriate range of values for θ. By looking at the graph we see that $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$. Our final equation is:

$$r = 0 \qquad \text{or} \qquad r = 2a\cos\theta, \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}.$$

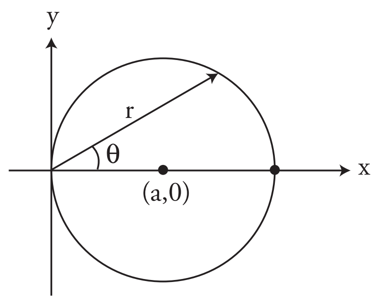


Figure 2: Off center circle in polar coordinates.

To check our work, let's find some points on this curve:

- At $\theta = 0$, $r = 2a$ and so $x = 2a$ and $y = 0$.
- At $\theta = \frac{\pi}{4}$, $r = 2a\cos\frac{\pi}{4} = a\sqrt{2}$. Hence $x = a$ and $y = a$.

Polar Coordinates and Area

How would we calculate an area using polar coordinates? Our basic increment of area will be shaped like a slice of pie. The slice of pie shown in Figure 1 has

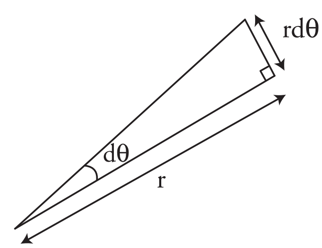


Figure 1: A slice of pie with radius r and angle $d\theta$.

a piece of a circular arc along its boundary with arc length $r\,d\theta$. We'll say that dA equals the area of the slice.

How do we express dA in terms of r and θ? The total area of the pie this was sliced from is πr^2. To find the area dA we note that the proportion of the total area covered equals the proportion of arc length covered. So:

$$\begin{aligned}\frac{dA}{\pi r^2} &= \frac{d\theta}{2\pi r}\\ dA &= \frac{r\,d\theta}{2\pi r}\cdot \pi r^2\\ dA &= \frac{1}{2}r^2\,d\theta\end{aligned}$$

This is the basic formula for an increment of area in polar coordinates.

We want to use polar coordinates to compute areas of shapes other than circles. In this case r will be a function of θ. The distance between the curve and the origin changes depending on what angle our ray is at. Our center point of reference is the origin; we think of rays emerging from the origin at some angle θ; $r(\theta)$ is, roughly, the distance we must travel along that ray to get to the curve.

To find the area of a shape like this, we break it up into circular sectors with angle $\Delta\theta$. Since the curve is not a circle the circular sectors won't perfectly cover the region, so we just approximate the area of a wedge between the curve and the origin by:

$$\Delta A \approx \frac{1}{2}r^2\Delta\theta.$$

If we take the limit as $\Delta\theta$ approaches zero our sum of sector areas will approach

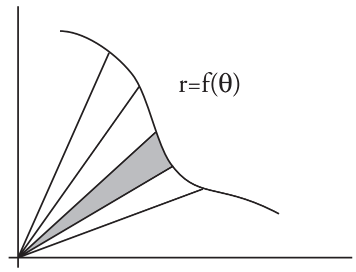


Figure 2: A slice from an oddly shaped pie.

the exact area and we get:

$$dA = \frac{1}{2} r^2 \, d\theta.$$

This is very similar to letting Δx go to zero in a Riemann sum of rectangle areas.

In the limit, we have:

$$A = \int_{\theta_1}^{\theta_2} \frac{1}{2} r^2 \, d\theta.$$

Remember that we're assuming r is a function of θ.

Area of an Off Center Circle

Let's find the area in polar coordinates of the region enclosed by the curve $r = 2a\cos\theta$. We've previously shown that this curve describes a circle with radius a centered at $(a, 0)$. In rectangular coordinates its equation is $(x-a)^2 + y^2 = a^2$.

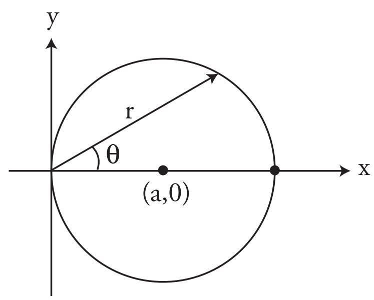


Figure 1: Off center circle $r = 2a\cos\theta$.

We're going to integrate an infintessimal amount of area dA. The integral will go from $\theta_1 = -\frac{\pi}{2}$ to $\theta_2 = \frac{\pi}{2}$. We could find these limits by looking at Figure 1; to draw the circle we might start by moving "down" at angle $-\frac{\pi}{2}$. As we move along the bottom of the circle toward $(2a, 0)$ the angle increases to 0, and as we trace out the top of the circle we're moving from angle 0 to angle $\frac{\pi}{2}$ ("up").

We might also find the limits of integration by looking at the formula and realizing that the cosine function is positive for $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$. When $\theta = \pm\frac{\pi}{2}$, $r = 2a\cos\theta$ is 0, so the two ends of the curve meet at the origin.

Our infinitessimal unit of area is $dA = \frac{1}{2}r^2\,d\theta$, so:

$$\begin{aligned}
A &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2}r^2\,d\theta \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2}(2a\cos\theta)^2\,d\theta \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2a^2\cos^2\theta\,d\theta \\
&= 2a^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\theta\,d\theta \qquad \text{(half angle formula)} \\
&= 2a^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1+\cos 2\theta}{2}\,d\theta \\
&= a^2\left[\theta + \frac{1}{2}\sin 2\theta\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}}
\end{aligned}$$

$$
\begin{aligned}
&= a^2\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) \\
&= \pi a^2.
\end{aligned}
$$

We know that the area of a circle of radius a is πa^2; our answer is correct.

Graph of $r = 2a\cos\theta$

Let's get some more practice in graphing and polar coordinates. We just found the area enclosed by the curve $r = 2a\cos\theta$ for $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$. What happens when θ doesn't lie in this range?

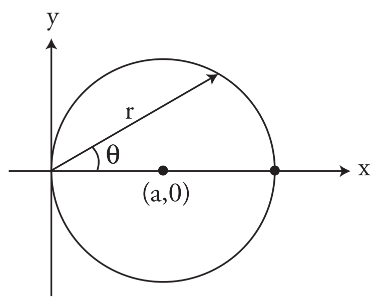


Figure 1: Off center circle $r = 2a\cos\theta$.

When $\frac{\pi}{2} < \theta < \pi$, r is negative. For example, when $\theta = \frac{3\pi}{4}$, $\cos\theta = -\frac{\sqrt{2}}{2}$ and $r = -a\sqrt{2}$. If we move a distance of *negative* $a\sqrt{2}$ in the direction of angle $\frac{3\pi}{4}$ we arrive at the point $(-a\sqrt{2}, \frac{3\pi}{4})$, which is $(a, -a)$ in rectangular coordinates.

In fact, because we know that the points on the curve must have the property:

$$(x-a)^2 + y^2 = a^2$$

in rectangular coordinates, we know that as θ increases, the point $(2a\cos\theta, \theta)$ must remain on that same curve. As θ ranges from 0 to 2π (or from $-\frac{\pi}{2}$ to $\frac{3\pi}{2}$), the point $(2a\cos\theta, \theta)$ travels around the circle twice.

A common mistake is to choose the wrong limits of integration and count the same area twice, or cancel a positive area with an overlapping negative one.

Question: Can you find the area using the limits of integration 0 and π?

Answer: Yes. The integral $\displaystyle\int_0^{\pi} \frac{1}{2}(2a\cos\theta)^2\, d\theta$ gives a correct answer.

However, $r = 2a\cos\theta$, $0 \le \theta \le \pi$ is an awkward way to describe a circle. As θ ranges from 0 to $\frac{\pi}{2}$, r is positive and (r, θ) moves along the top half of the circle. As θ sweeps through the second quadrant ($\frac{\pi}{2} < \theta < \pi$), r is negative and so the curve appears in the fourth quadrant.

When we work with negative values of r it's easy to get confused, so when possible it's a good idea to choose our limits of integration so that r is positive.

Graph of $r = \sin 2\theta$

This curve is a favorite; a similar curve appears in the homework. We'll plot a few points $(\sin 2\theta, \theta)$ to get an idea of what the graph of this curve looks like.

θ	$r = \sin 2\theta$
0	0
$\frac{pi}{4}$	1
$\frac{pi}{2}$	0

Note that $\sin 2\theta > 0$ for $0 < \theta < \frac{\pi}{2}$. So the curve starts at the origin, goes outward to the point $(1, \frac{\pi}{4})$ (which is $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ in rectangular coordinates), then returns to the origin as θ moves to point "up".

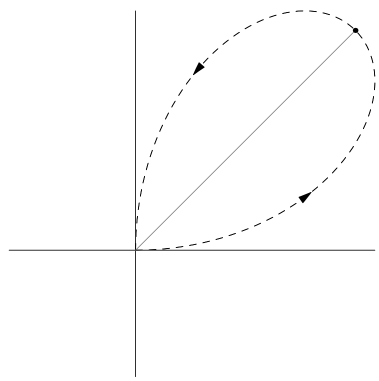


Figure 1: Graph of $r = \sin 2\theta$ for $0 < \theta < \frac{\pi}{2}$.

Because of the symmetries of the sine function, the curve will do something similar in each quadrant. However, it's useful to watch the curve being drawn in order to understand how its parts are connected.

The "loops" in the graph are caused by $r = \sin 2\theta$ changing sign each time the graph intersects the origin. When $\frac{\pi}{2} < \theta < \pi$, our angle is in the second quadrant; the portion of the graph corresponding to those values of θ appears opposite the angle, in the fourth quadrant.

If we want to compute the area of one petal of this rose, we have to be careful to use the right bounds for θ.

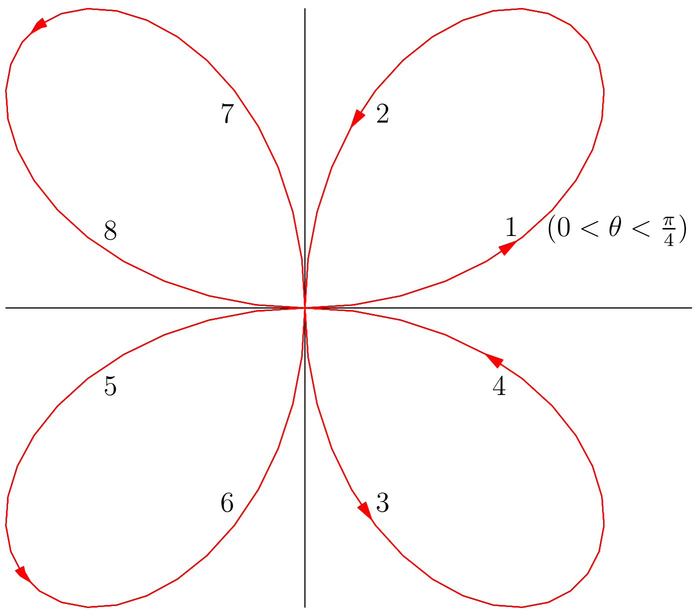


Figure 2: Graph of $r = \sin 2\theta$; a four leaf rose.

Polar Coordinates and Conic Sections

Suppose we want to graph the curve described by:

$$r = \frac{1}{1 + 2\cos\theta}.$$

Again we start by plotting some points on this curve:

θ	r
0	$\frac{1}{3}$
$\frac{\pi}{2}$	-1
$-\frac{\pi}{2}$	1

By using the equations:

$$x = r\cos\theta, \qquad y = r\sin\theta$$

we can convert these polar coordinates to rectangular coordinates, show in Figure 1. For example, when $\theta = \frac{\pi}{2}$ we know that $r = 1$ and so:

$$\begin{aligned} x = r\cos\theta &= 0 \\ y = r\sin\theta &= -1 \end{aligned}$$

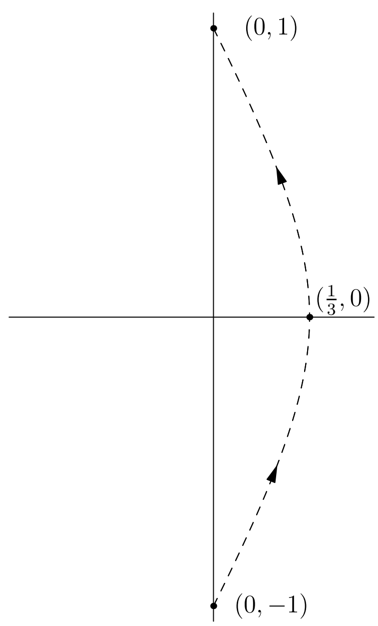


Figure 1: Rectangular coordinates of points on the curve $r = \dfrac{1}{1 + 2\cos\theta}$.

When $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$, the denominator $1 + 2\cos\theta$ is positive and so r is positive; the curve traced over this interval must look something like the dotted line in Figure 1.

It's possible for the denominator to be 0:

$$\begin{aligned}
1 + 2\cos\theta &= 0 \\
2\cos\theta &= -1 \\
\cos\theta &= -\frac{1}{2} \\
\theta &= \arccos\left(-\frac{1}{2}\right) \\
\theta &= \pm\frac{2\pi}{3}
\end{aligned}$$

The radius r goes to infinity as θ approaches $\frac{2\pi}{3}$ and $-\frac{2\pi}{3}$, so the curve will extend infinitely far in those directions.

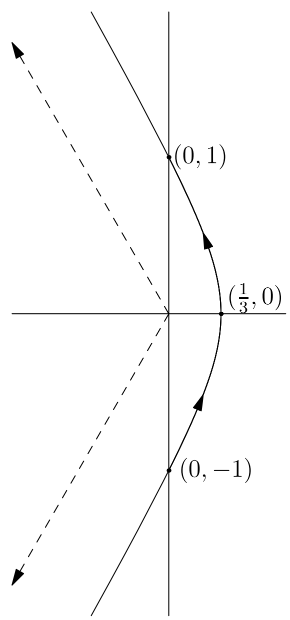


Figure 2: Graph of the curve $r = \dfrac{1}{1 + 2\cos\theta}$.

This is as much as we'll be able to figure out about the graph without converting its equation from polar to rectangular coordinates. Let's do that.

Rectangular Equation

What is the rectangular $(x,\, y)$ equation for $r = \dfrac{1}{1 + 2\cos\theta}$?

To answer this question we could use our formula $x = r\cos\theta$, $y = r\sin\theta$ and then try to simplify, but if we're clever we can manipulate our original formula until it appears in terms of x and y.

$$\begin{aligned} r &= \frac{1}{1+2\cos\theta} \\ r + 2r\cos\theta &= 1 \\ r &= 1 - 2r\cos\theta \\ r &= 1 - 2x \end{aligned}$$

Multiplying both sides by the denominator simplified the expression and allowed us to make the substitution $x = r\cos\theta$. The variable θ no longer appears.

If we square both sides of this new equation we can get rid of the variable r as well:

$$\begin{aligned} r &= 1 - 2x \\ r^2 &= (1-2x)^2 \\ x^2 + y^2 &= 1 - 4x + 4x^2 \\ -3x^2 + y^2 + 4x - 1 &= 0 \end{aligned}$$

This is a standard calculation with a standard result; whenever we start with have $\dfrac{1}{a + b\cos\theta}$ or $\dfrac{1}{a + b\sin\theta}$ we'll end up with an equation like this.

Because the signs of the coefficients of x^2 and y^2 are different, this is the equation of a hyperbola. (If the signs match, the equation describes an ellipse; if one of these coefficients is 0 the graph is a parabola.) We can now conclude that the dotted lines in Figure 1 are asymptotes of the graph.

To complete our understanding of the curve $r = \frac{1}{1+2\cos\theta}$ we ask what happens when the denominator $1 + 2\cos\theta$ is negative?

Since we know that the equation for the curve in rectangular coordinates is $-3x^2 + y^2 + 4x - 1 = 0$, we can guess that for $\frac{3\pi}{2} < \theta < \frac{5\pi}{2}$ the curve must trace out the other branch of the hyperbola.

Connection to Kepler's Second Law

There is a beautiful connection between the basic formula for area and these types of curve.

As you may know, the trajectories of comets are hyperbolas. Ellipses are the trajectories of planets or asteroids. When you represent hyperbolas and ellipses in polar coordinates like this, it turns out that:

$$r = 0 \qquad \text{is the focus of the hyperbola.}$$

Polar coordinates are the natural way to express the trajectory of a planet or comet if you want the center of gravity (the sun) to be at the origin.

The formula
$$dA = \frac{1}{2} r^2 \, d\theta$$

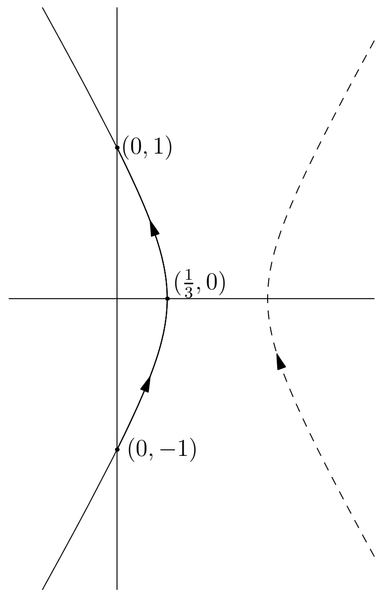


Figure 3: Both branches of the curve $r = \dfrac{1}{1+2\cos\theta}$.

is a central formula in astronomy. Kepler's second law says that a line joining a comet or planet to the sun sweeps out equal areas in equal time periods. In other words, the rate of change of area swept out is constant:

$$\frac{dA}{dt} = \text{constant}.$$

This tells us that as a comet travels around the sun, its speed varies, and varies predictably. Since we know $dA = \frac{1}{2}r^2\,d\theta$, we can conclude that:

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt}.$$

Combining this with Kepler's second law we get:

$$\frac{1}{2}r^2\frac{d\theta}{dt} = \text{constant}.$$

In modern-day terms, what this formula says is that angular momentum is conserved. The objects Kepler was observing weren't subject to friction or air resistance, but this equation is the same one used to describe why a top keeps spinning after you let it go, or why an ice skater spins faster when she pulls her arms and legs in.

Review for Test 4

Exam 4 will test the following concepts:

1. Techniques of Integration (55%)
    - Trig substitution and trig integrals.
    - Integration by parts.
    - Partial fractions.
2. Parametric Curves (45%)
    - Arc length.
    - Area of surfaces of revolution.
    - Polar coordinates, including area.

There are six problems on the test; they are somewhat similar to the questions on the sample test.

Student Questions on Test 4

Question: Did we do arc length in polar coordinates?

Answer: No, we did not. We'd need to know about arc length in polar coordinates in order to, for example, predict the speed of a comet, but that topic won't be on this exam.

Question: Will be asked to sketch graphs of curves like $r = \sin 3\theta$?

Answer: For a problem this complicated there are two possibilities. Either we take a long time to sketch it out or we get a hint that the curve traces out a three leafed rose.

In general, we won't be told which techniques of integration to use to compute the integrals on the test. However, for integrals we're likely to get stuck on, we'll get a hint or instructions on how to proceed (or not to proceed).

In setting up integrals from the later half of the unit we'll always need to do three key steps: finding the lower limit, the upper limit, and the integrand. The second step would be to evaluate the integral, which may or may not be required on the test (or possible).

Question: Can you review what to do when the denominator of a partial fractions problem has a repeated factor?

Answer: Suppose we have:

$$\frac{x^2+21}{(x+2)^2\,x\,(x+1)}$$

We'll have one variable for each factor in the denominator. The setup looks like:

$$\frac{x^2+21}{(x+2)^2\,x\,(x+1)} = \frac{A}{x+2} + \frac{B}{(x+2)^2} + \frac{C}{x} + \frac{D}{x+1}.$$

If we instead have

$$\frac{x^2+21}{(x+2)^3\,x\,(x+1)},$$

the setup will look like:

$$\frac{x^2+21}{(x+2)^3\,x\,(x+1)} = \frac{A}{x+2} + \frac{B}{(x+2)^2} + \frac{E}{(x+2)^3} + \frac{C}{x} + \frac{D}{x+1}.$$

The more repeated roots there are, the harder the problem gets. The cover up method only helps you solve for one variable for each different factor in the denominator; we can use it to solve for C, D and E but not for A or B. To find A and B you'll either have to plug in values or use other algebraic tools for solving a system of equations.

Question: Does the $x^3 + 21$ in the numerator affect the setup?

Answer: The answer is almost "no". The setup will always look like this, but if the degree of the numerator is too large you'll have something like an improper fraction and will need to use long division before you can apply the method of partial fractions.

Question: Will we need to know how to do reduction formulas?

Answer: If a reduction formula or something else out of the ordinary is required to complete an integral, you will be told what you need to do.

Question: In the partial fractions method, what happens if you have a quadratic in the denominator?

Answer: Suppose we're asked to decompose:

$$\frac{x^2+21}{(x^2+2)^2\,x\,(x+1)}.$$

Our setup would look like:

$$\frac{x^2+21}{(x^2+2)^2\,x\,(x+1)} = \frac{A_1x+B_1}{x^2+2} + \frac{A_2x+B_2}{(x^2+2)^2} + \frac{C}{x} + \frac{D}{x+1}.$$

Example: $\int x \tan^{-1}(x)\, dx$

This is a slightly harder problem than any on the test, but something like this might appear on the final. How should we approach this?

Student: Integration by parts.

Great! Because $\tan^{-1}(x)$ is begging to be differentiated to be made simpler. So we choose:

$$u = \tan -1(x), \qquad v' = x,$$
$$u' = \frac{1}{1+x^2}, \qquad v = \frac{x^2}{2}.$$

Integration by parts then gives us:

$$\int x \tan^{-1}(x)\, dx = \frac{x^2}{2} \tan^{-1}(x) - \int \frac{x^2}{2} \cdot \frac{1}{1+x^2}\, dx$$

We're not done yet! We still have to integrate:

$$-\frac{1}{2} \int \frac{x^2}{1+x^2}\, dx.$$

What do we do know?

Student: Trig substitution.

Trig substitution will work, but that's not what I had in mind.

Student: Add and subtract one in the numerator.

That's a good idea. This is an example of a rational expression in which the numerator and denominator have the same degree, so you could use long division to simplify the "improper fraction". An equivalent shortcut is:

$$\begin{aligned}
\frac{x^2}{1+x^2} &= \frac{x^2+1-1}{1+x^2} \\
&= \frac{x^2+1}{1+x^2} - \frac{1}{1+x^2} \\
&= 1 - \frac{1}{1+x^2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
-\frac{1}{2} \int \frac{x^2}{1+x^2}\, dx &= -\frac{1}{2} \int \left(1 - \frac{1}{1+x^2}\right) dx \\
&= -\frac{1}{2} x + \frac{1}{2} \tan^{-1} x + C.
\end{aligned}$$

The answer to the original question is then:

$$\begin{aligned}
\int x \tan^{-1}(x)\, dx &= \frac{x^2}{2} \tan^{-1}(x) - \int \frac{x^2}{2} \frac{1}{1+x^2}\, dx \\
&= \frac{x^2}{2} \tan^{-1}(x) + \frac{1}{2} x - \frac{1}{2} \tan^{-1} x + c.
\end{aligned}$$

Introduction to L'Hôpital's Rule

In this final unit we tie up some loose ends related to calculus and limits. Our first topic is L'Hôpital's rule, which is useful for understanding multiplication and division by infinity.

L'Hôpital's rule is also known as L'Hospital's rule; the circumflex accent indicates that the letter S has been omitted, so the two spellings are equivalent. The two spellings are pronounced identically, with a long O and silent S.

L'Hôpital's rule is used to calculate limits of expressions like:

$$\begin{aligned} & x \ln x && \text{as } x \to 0^+, \\ & xe^{-x} && \text{as } x \to \infty, \\ & \frac{\ln x}{x} && \text{as } x \to \infty. \end{aligned}$$

We could use a calculator to guess what these limits are, but L'Hôpital's rule gives us a systematic and provable way of finding the limits.

Elementary Example of L'Hôpital's Rule

We begin by applying L'Hôpital's rule to a problem we could have solved earlier:

$$\lim_{x \to 1} \frac{x^{10} - 1}{x^2 - 1}.$$

We listed some categories of limits at the beginning of the course; this falls into the category of "interesting limits" because if we just plug in $x = 1$ we get $\frac{0}{0}$. This is called an *indeterminate form.*

To find the limit using techniques we already know, we'd do the following:

$$\lim_{x \to 1} \frac{x^{10} - 1}{x^2 - 1} = \lim_{x \to 1} \frac{(x^{10} - 1)/(x - 1)}{(x^2 - 1)/(x - 1)}.$$

We could calculate $(x^{10} - 1)/(x - 1)$ using long division, but that's a long calculation. We can find this limit more quickly using calculus.

We've used calculus to understand a fraction in indeterminate form when we studied the difference quotient. If $f(x) = x^{10} - 1$, then $f(1) = 0$ and the difference quotient is:

$$\frac{f(x) - f(1)}{(x - 1)} = \frac{x^{10} - 1}{x - 1}.$$

We know from our studies of difference quotients that:

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

We conclude that:

$$\lim_{x \to 1} \frac{x^{10} - 1}{x - 1} = f'(1) = 10.$$

Our expression:

$$\frac{x^{10} - 1}{x^2 - 1} = \frac{(x^{10} - 1)/(x - 1)}{(x^2 - 1)/(x - 1)}$$

describes a ratio of difference quotients, so if $g(x) = x^2 - 1$ this line of reasoning tells us that:

$$\begin{aligned}
\lim_{x \to 1} \frac{x^{10} - 1}{x^2 - 1} &= \lim_{x \to 1} \frac{(x^{10} - 1)/(x - 1)}{(x^2 - 1)/(x - 1)} \\
&= \frac{\lim_{x \to 1}((x^{10} - 1)/(x - 1))}{\lim_{x \to 1}((x^2 - 1)/(x - 1))} \\
&= \frac{f'(1)}{g'(1)} \\
&= \frac{10}{2} \\
&= 5.
\end{aligned}$$

Dividing by $x - 1$ and interpreting the fraction as a ratio of difference quotients enabled us to solve the problem by taking two easy derivatives and saved us from a lengthy exercise in long division.

Why L'Hôpital's Rule Works

Suppose we're considering a limit:

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

that is indeterminate; i.e. $f(a) = g(a) = 0$.

We do the same thing we did in the previous example:

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

We can use this formula to calculate limits of indeterminate expressions provided that $g'(a) \neq 0$.

Question: Is there a more intuitive way to understand this rule?

Answer: There are other ways to understand the rule, but none that are much more intuitive than this. It helps to understand these other ways as well. For example, we'll soon see how we can derive l'Hôpital's rule by looking at the linearizations of f and g at a.

L'Hôpital's Rule, Version 1

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

provided $f(a) = g(a) = 0$ and the right hand limit exists.

This version of l'Hôpital's rule is similar to but better than the rule we just derived; it lets us get rid of the restriction $g'(a) \neq 0$.

This is practically the same thing we did in computing $\lim_{x \to 1} \frac{x^{10} - 1}{x^2 - 1}$. First we took the derivative of the numerator and denominator separately (do *not* apply the quotient rule here!)

$$\lim_{x \to 1} \frac{x^{10} - 1}{x^2 - 1} = \lim_{x \to 1} \frac{10x^9}{2x}.$$

Next, we found the limit as x approaches 1:

$$\lim_{x \to 1} \frac{10x^9}{2x} = \frac{10}{2} = 5.$$

Instead of evaluating $\frac{x^{10} - 1}{x^2 - 1}$, we took derivatives and instead evaluated $\frac{10x^9}{2x}$ which is much simpler. This is typical of l'Hôpital's rule.

Example: $\lim_{x \to 0} \dfrac{\sin 5x}{\sin 2x}$

This is similar to an example we saw earlier in the course. Here $f(x) = \sin 5x$, $g(x) = \sin 2x$, and $a = 0$. Since $f(a) = g(a) = \sin 0 = 0$, we can apply l'Hôpital's rule and find this limit:

$$\begin{aligned} \lim_{x \to 0} \frac{\sin 5x}{\sin 2x} &= \lim_{x \to 0} \frac{5 \cos 5x}{2 \cos 2x} \qquad \text{(l'Hop)} \\ &= \lim_{x \to 0} \frac{5 \cos(5 \cdot 0)}{2 \cos(2 \cdot 0)} \\ &= \frac{5}{2}. \end{aligned}$$

Repeating L'Hôpital's Rule

This example illustrates the superiority of Version 1 of l'Hôpital's rule; it works even if $g'(a) = 0$.

In this case, $f(x) = \cos x - 1$, $g(x) = x^2$, and $a = 0$. We're trying to find:

$$\lim_{x \to 0} \frac{\cos x - 1}{x^2}.$$

We can easily verify that $f(a) = g(a) = 0$.

We apply l'Hôpital's rule:

$$\lim_{x \to 0} \frac{\cos x - 1}{x^2} = \lim_{x \to 0} \frac{-\sin x}{2x}.$$

Notice that $\dfrac{-\sin x}{2x}$ is undefined when $x = 0$; it's of the type $\frac{0}{0}$. That's OK; this version of l'Hôpital's rule still works when $g'(x) = 0$; it works as long as $\lim_{x \to a} \frac{f'(a)}{g'(a)}$ is defined.

We need to find $\displaystyle\lim_{x \to 0} \frac{-\sin x}{2x}$. We can do that by applying l'Hôpital's rule!

$$\lim_{x \to 0} \frac{-\sin x}{2x} = \lim_{x \to 0} \frac{-\cos x}{2}.$$

All together, the calculation looks like:

$$\begin{aligned}
\lim_{x \to 0} \frac{\cos x - 1}{x^2} &= \lim_{x \to 0} \frac{-\sin x}{2x} && \text{(l'Hop)} \\
&= \lim_{x \to 0} \frac{-\cos x}{2} && \text{(l'Hop)} \\
&= \frac{-\cos 0}{2} \\
&= -\frac{1}{2}.
\end{aligned}$$

Notice that we only know that the hypotheses of l'Hôpital's rule hold for our problem when we reach the end and get the limit. The theorem says that l'Hôpital's rule only works if the limit exists, and we find out that the limit exists by using l'Hôpital's rule to calculate it. This is logically somewhat subtle, but in practice the theorem works very well.

Question: Why does the limit have to exist? Isn't it just the derivative that has to exist?

Answer: No, we need the derivative of the numerator, the derivative of the denominator *and* the limit to exist.

Surprisingly enough, we don't need $f'(a)$ to exist; we're working with limits as x approaches a, so what happens when $x = a$ doesn't necessarily matter to us. Once again we're using limits to get close to something that's not defined at the exact value $x = a$.

If any one of the three limits $\lim_{x\to a} f'(x)$, $\lim_{x\to a} g'(x)$, and $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ does not exist then we can't apply l'Hôpital's rule. This is because the theorem we used to prove the rule works might not be true if these limits are undefined.

To get a little ahead of ourselves, notice that in:

$$\lim_{x\to\infty} xe^{-x}$$

the limit $\lim_{x\to\infty} x$ is undefined. Nevertheless, l'Hôpital's rule will apply here.

Comparison With Approximation

We just used l'Hôpital's rule to show that:

$$\lim_{x\to 0} \frac{\cos x - 1}{x^2} = -\frac{1}{2}.$$

Let's compare that to what we get using the method of approximations, replacing the functions involved with their linear or quadratic approximations.

In the example of $\lim_{x\to 0} \frac{\sin 5x}{\sin 2x}$, we would use the linear approximation $\sin u \approx u$ for u near 0 to get:

$$\frac{\sin 5x}{\sin 2x} \approx \frac{5x}{2x} = \frac{5}{2}$$

for x near 0. Since the approximation $\sin u \approx u$ becomes exact as u approaches 0, we could then conclude that:

$$\lim_{x\to 0} \frac{\sin 5x}{\sin 2x} = \frac{5}{2}.$$

To use this method to find:

$$\lim_{x\to 0} \frac{\cos x - 1}{x^2},$$

we would approximate:

$$\frac{\cos x - 1}{x^2}$$

by:

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

Again the approximation becomes exact as x approaches 0, so:

$$\lim_{x\to 0} \frac{\cos x - 1}{x^2} = -\frac{1}{2}.$$

The method of approximation gives us the same result as l'Hôpital's rule, as it should. Both methods are valid and they both involve about the same amount of work. The approximation we used for the cosine function is related to the second derivative of $\cos x$, and we had to find the second derivative of $\cos x$ when we applied l'Hôpital's rule.

Extensions of L'Hôpital's Rule

Our first version of l'Hôpital's rule told us that:

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

provided that $f(a) = g(a) = 0$ and $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists.

We've seen that we can get nearly equivalent results by replacing $f(x)$ and $g(x)$ by linear or quadratic approximations. L'Hôpital's rule is superior to the method of approximation because it works better in some situations.

It turns out that l'Hôpital's rule works even under the following conditions:

- $a = \pm\infty$
- $f(a), g(a) = \pm\infty$
- $\lim_{x\to a} \frac{f'(a)}{g'(a)} = \pm\infty$

In other words, l'Hôpital's rule works not just in the $\frac{0}{0}$ case but also when you're taking a limit of the form $\frac{\infty}{\infty}$. It will give us the right answer if $\frac{f'(a)}{g'(a)}$ approaches $-\infty$, ∞ or some finite number. It fails if $\frac{f'(a)}{g'(a)}$ oscillates wildly, but we don't encounter those conditions in this class; l'Hôpital's rule handles everything we could expect it to handle, and it's easy to use.

Rate of Growth of $\ln x$

This expression is in indeterminate form but looks like it might be the wrong type. This isn't a fraction, so we have to think about how to apply l'Hôpital's rule.

In the expression, the factor x is approaching 0 while the factor $\ln x$ is approaching negative infinity.

$$\lim_{x \to 0^+} \underbrace{x}_{\to 0} \underbrace{\ln x}_{\to -\infty}$$

We're multiplying a number that's getting smaller and smaller by one that's getting larger and larger; the result could be really large or really small, depending on rates of growth.

The first step in finding the limit is to rewrite the expression as a ratio, rather than as a product. We'll choose to write it as:

$$x \ln x = \frac{\ln x}{1/x}.$$

This is an expression of the type $\frac{-\infty}{\infty}$, which is one of the forms we can apply l'Hôpital's rule to. Let's do that:

$$\begin{aligned} \lim_{x \to 0^+} x \ln x &= \lim_{x \to 0^+} \frac{\ln x}{1/x} \\ &= \lim_{x \to 0^+} \frac{1/x}{-1/x^2} \qquad \text{(l'Hop)} \\ &= \lim_{x \to 0^+} -x \\ &= 0. \end{aligned}$$

We conclude that x goes to 0 faster than $\ln x$ goes to negative infinity, and so the limit of the product is 0. You might not have been able to guess this in advance.

Rate of Growth of e^{px}

When we looked at $\lim_{x\to 0^+} x \ln x$ we found that the value of the limit was 0, so x shrinks to 0 faster than $\ln x$ grows to negative infinity. The next two examples illustrate similar rate properties, which will be important when we study improper integrals and elsewhere.

Example: $\displaystyle\lim_{x\to\infty} xe^{-px}, \quad (p > 0)$

The expression xe^{-px} is a product, not a ratio, so we need to rewrite it before we use l'Hôpital's rule. We choose to rewrite it as $\dfrac{x}{e^{px}}$. This is of the form $\frac{\infty}{\infty}$, so we can use l'Hôpital's rule to calculate:

$$\begin{aligned}
\lim_{x\to\infty} xe^{-px} &= \lim_{x\to\infty} \frac{x}{e^{px}} \\
&= \lim_{x\to\infty} \frac{1}{pe^{px}} \qquad \text{(l'Hop)} \\
&= \frac{1}{\infty} \\
&= 0.
\end{aligned}$$

We conclude that when $p > 0$, x grows more slowly than e^{px} as x goes to infinity.

Example: $\displaystyle\lim_{x\to\infty} \frac{e^{px}}{x^{100}} \quad (p > 0)$

This example doesn't give us much more information, but it's good practice. The value of this limit gives us information about the relative rates of growth of e^{px} and x^{100}.

The expression $\displaystyle\lim_{x\to\infty} \frac{e^{px}}{x^{100}}$ is of the form $\frac{\infty}{\infty}$, so we can use l'Hôpital's rule again. In fact, there are two ways we could use l'Hôpital's rule. The slow way looks like:

$$\begin{aligned}
\lim_{x\to\infty} \frac{e^{px}}{x^{100}} &= \lim_{x\to\infty} \frac{pe^{px}}{100x^{99}} \qquad \text{(l'Hop)} \\
&= \lim_{x\to\infty} \frac{p^2 e^{px}}{100 \cdot 99 x^{98}} \qquad \text{(l'Hop)} \\
&= \lim_{x\to\infty} \frac{p^3 e^{px}}{100 \cdot 99 \cdot 98 x^{97}} \qquad \text{(l'Hop)} \\
&\vdots
\end{aligned}$$

We could apply l'Hôpital's rule 100 times and we'd eventually get an answer.

The clever way is to rewrite the expression as follows:

$$\lim_{x\to\infty} \frac{e^{px}}{x^{100}} = \left(\lim_{x\to\infty} \frac{e^{px/100}}{x}\right)^{100}$$

$$
\begin{aligned}
&= \left(\lim_{x\to\infty} \frac{\frac{p}{100}e^{px/100}}{1}\right)^{100} \qquad \text{(l'Hop)}\\
&= \left(\lim_{x\to\infty} \frac{p\cdot e^{px/100}}{100}\right)^{100}\\
&= \infty
\end{aligned}
$$

In this example $\lim_{x\to a} \frac{f'(a)}{g'(a)} = \infty$, another possible outcome of l'Hôpital's rule. We conclude that e^{px} grows faster than x^{100} when p is positive. In fact, e^{px} grows faster than *any* polynomial in x; exponential functions grow faster than powers of x.

Comparing Growth of $ln(x)$ and $x^{\frac{1}{3}}$

We have one more item on our original list of limits to cover; again we'll look at a slight variation on the original problem. We're going to find:

$$\lim_{x\to\infty} \frac{\ln x}{x^{1/3}}.$$

This limit is of the form $\frac{\infty}{\infty}$, so we apply l'Hôpital's rule to find:

$$\begin{aligned}
\lim_{x\to\infty} \frac{\ln x}{x^{1/3}} &= \lim_{x\to\infty} \frac{1/x}{\frac{1}{3}x^{-2/3}} \qquad \text{(l'Hop)} \\
&= \lim_{x\to\infty} 3x^{-1/3} \\
&= 0
\end{aligned}$$

We conclude that $\ln x$ grows more slowly as x approaches infinity than $x^{1/3}$ or any positive power of x. In other words, $\ln x$ increases very slowly.

Question: When we discussed extensions of l'Hôpital's rule, we learned that we're allowed to change some hypotheses. How many hypotheses can we change at once?

Answer: We can make any or all of the three changes listed. However, $\dfrac{f(a)}{g(a)}$ must always be of the form $\frac{\infty}{\infty}$, $-\frac{\infty}{\infty}$, or $\frac{0}{0}$.

The Indeterminate Form 0^0

We next consider the limit:

$$\lim_{x\to 0^+} x^x.$$

Can we compute this?

There are many different indeterminate forms; x^x is one of the simpler examples. In this case, because x is a moving exponent, we can use a trick to evaluate the limit.

Since we have a moving exponent, we will use base e. We rewrite our original expression as follows:

$$x^x = e^{x \ln x}.$$

Now we can focus our attention on the exponent:

$$\begin{aligned}
\lim_{x\to 0^+} x \ln x &= \lim_{x\to 0^+} \frac{\ln x}{1/x} \\
&= \lim_{x\to 0^+} \frac{1/x}{-1/x^2} \qquad \text{(l'Hop)} \\
&= \lim_{x\to 0^+} -x \\
&= 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\lim_{x\to 0^+} x^x &= \lim_{x\to 0^+} e^{x \ln x} \\
&= e^0 \\
&= 1.
\end{aligned}$$

This was relatively easy to calculate because we have so many powerful tools to work with.

Example: $\lim_{x\to 0} \frac{\sin x}{x^2}$

If we apply l'Hôpital's rule to this problem we get:

$$\begin{aligned}\lim_{x\to 0} \frac{\sin x}{x^2} &= \lim_{x\to 0} \frac{\cos x}{2x} \qquad \text{(l'Hop)}\\ &= \lim_{x\to 0} \frac{-\sin x}{2} \qquad \text{(l'Hop)}\\ &= 0.\end{aligned}$$

If we instead apply the linear approximation method and plug in $\sin x \approx x$, we get:

$$\begin{aligned}\frac{\sin x}{x^2} &\approx \frac{x}{x^2}\\ &\approx \frac{1}{x}.\end{aligned}$$

We then conclude that:

$$\begin{aligned}\lim_{x\to 0^+} \frac{\sin x}{x^2} &= \infty\\ \lim_{x\to 0^-} \frac{\sin x}{x^2} &= -\infty.\end{aligned}$$

There's something fishy going on here. What's wrong?

Student: L'Hôpital's rule wasn't applied correctly the second time.

That's correct; $\lim_{x\to 0} \frac{\cos x}{2x}$ is of the form $\frac{1}{0}$, not $\frac{0}{0}$ or some other indeterminate form.

This is where you have to be careful when using l'Hôpital's rule. You have to verify that you have an indeterminate form like $\frac{0}{0}$ or $\frac{\infty}{\infty}$ before applying the rule. The moral of the story is: **Look before you l'Hôp**.

Also, don't use l'Hospital's rule as a crutch. If we want to evaluate:

$$\lim_{x\to\infty} \frac{x^5 - 2x^4 + 1}{x^4 + 2}$$

we can apply l'Hôpital's rule four times, or we could divide the numerator and denominator by x^5 to conclude:

$$\begin{aligned}\lim_{x\to\infty} \frac{x^5 - 2x^4 + 1}{x^4 + 2} &= \lim_{x\to\infty} \frac{1 - 2/x + 1/x^5}{1/x + 2/x^5}\\ &= \frac{1}{0}\\ &= \infty.\end{aligned}$$

After enough practice with rates of growth, we can calculate this limit almost instantly:

$$\lim_{x\to\infty} \frac{x^5 - 2x^4 + 1}{x^4 + 2} \sim \lim_{x\to\infty} \frac{x^5}{x^4} = \infty.$$

l'Hôpital's Rule, Continued

In keeping with the spirit of "dealing with infinity" we look at an application of l'Hôpital's rule to a limit of the form $\frac{\infty}{\infty}$. In other words, as x approaches a we have:

- $f(x) \to \infty$
- $g(x) \to \infty$
- $\dfrac{f'(x)}{g'(x)} \to L$

and so we can conclude that:

$$\lim_{x \to a} \frac{f(x)}{g(x)} = L.$$

(Recall that a and L may be infinite.)

Rates of Growth

We apply this to "rates of growth"; the study of how rapidly functions increase. We know that the functions $\ln x$ and x^2 both go to infinity as x goes to infinity, and that x^2 increases much more rapidly than $\ln x$. We can formalize this idea as follows:

If $f(x) > 0$ and $g(x) > 0$ as x approaches infinity, then

$$f(x) << g(x) \,\text{as}\, x \to \infty \quad \text{means} \quad \lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

(Read $f(x) << g(x)$ as "$f(x)$ is a lot less than $g(x)$".) In our example, $f(x) = \ln x$ and $g(x) = x^2$. If we use l'Hôpital's rule to evaluate $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{\ln x}{x^2}$ we get:

$$\begin{aligned} \lim_{x\to\infty} \frac{\ln x}{x^2} &= \lim_{x\to\infty} \frac{\frac{1}{x}}{2x} \\ &= \lim_{x\to\infty} \frac{1}{2x^2} \\ &= 0. \end{aligned}$$

We conclude that $\ln x << x^2 \,\text{as}\, x \to \infty$.

If $p > 0$ then:

$$\ln x << x^p << e^x << e^{x^2} \,\text{as}\, x \to \infty.$$

Rates of Decay

"Rates of decay" are rates at which functions tend to 0 as x goes to infinity. Again our new notation comes in handy; if $p > 0$ then:

$$\frac{1}{\ln x} >> \frac{1}{x^p} >> e^{-x} >> e^{-x^2} \,\text{as}\, x \to \infty.$$

Introduction to Improper Integrals

An *improper integral* of a function $f(x) > 0$ is:

$$\int_a^{\infty} f(x)\,dx = \lim_{N\to\infty} \int_a^N f(x)\,dx.$$

We say the improper integral *converges* if this limit exists and *diverges* otherwise.

Geometrically then the improper integral represents the total area under a curve stretching to infinity. If the integral $\int_a^{\infty} f(x)\,dx$ converges the total area under the curve is finite; otherwise it's infinite. (See Figure 1.)

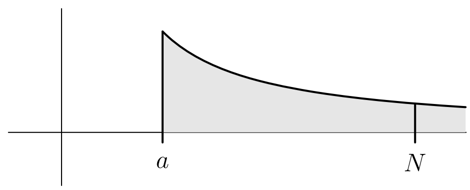


Figure 1: Infinite area under a curve.

How can an area that extends to infinity be finite? Obviously the area between a and N (i.e. $\int_a^N f(x)\,dx$) is finite. As N goes to infinity this quantity will either grow without bound or it will converge to some finite value. Our next step is to look at examples of each of these possibilities.

Example: $\displaystyle\int_0^\infty e^{-kx}\,dx \quad (k > 0)$

This is the most fundamental, by far, of the improper integrals. We start by calculating $\int_0^N e^{-kx}\,dx$:

$$\begin{aligned}
\int_0^N e^{-kx}\,dx &= -\frac{1}{k}e^{-kx}\Big|_0^N \\
&= -\frac{1}{k}e^{-kN} - -\frac{1}{k}e^0 \\
&= -\frac{1}{k}e^{-kN} + \frac{1}{k}
\end{aligned}$$

As N goes to infinity the $\frac{1}{k}$ does not change, but $-\frac{1}{k}e^{-kN}$ gets closer and closer to zero. (This is only true if k is positive!) So

$$\int_0^\infty e^{-kx}\,dx = \lim_{N\to\infty}\int_0^N e^{-kx}\,dx = \frac{1}{k}.$$

We can abbreviate this calculation as follows:

$$\begin{aligned}
\int_0^\infty e^{-kx}\,dx &= -\frac{1}{k}e^{-kx}\Big|_0^\infty \\
&= -\frac{1}{k}e^{-\infty} - -\frac{1}{k}e^0 \quad \text{(using the fact that } k > 0\text{)} \\
&= -0 + \frac{1}{k} \\
&= \frac{1}{k}
\end{aligned}$$

Question: What if the limit is infinity?

Answer: Good question. There's a difference between the limit existing and the limit being inifinte. Where our definition of improper integral says "if this limit exists" it means "exists and is finite"; if infinite limits are allowed they're mentioned explicitly, as in l'Hôpital's rule.

There is another part of this subject which we will not study here. If f changes sign (e.g. $f(x) = \frac{\sin x}{x}$) there can be some cancellation in the integral as f oscillates. Sometimes the limit exists, but the total area enclosed above and below the x-axis is infinite. In order to avoid this possibility we require that $f(x) > 0$.

Physical Interpretation

The number of radioactive particles in some radioactive substance that decay in time $0 \le t \le T$ is given (on average) by:

$$\int_0^T Ae^{-kt}\,dt.$$

If we let T go to infinity we get:

$$\int_0^\infty Ae^{-kt}\,dt = \frac{A}{k} = \text{ total number of particles.}$$

The notion that T goes to infinity is an idealization; we're not actually going to wait forever for the substance to decay. However, it's useful for us to write down and use this quantity even if it's not physically realistic. One reason to use this value as opposed to those for finite time intervals is that $\frac{A}{k}$ is much easier to work with than $-\frac{A}{k}e^{-kT} + \frac{A}{k}$.

Example: $\int_{-\infty}^{\infty} e^{-x^2}\,dx$

Another famous improper integral is:

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

This is a key number in probability. It is used to understand standard deviation, predict outcomes of elections, calculate insurance rates, and estimate income from lotteries.

This number was first calculated numerically around the end of the the seventeenth century by de Moivre, who was selling his services to various royalty who were running lotteries. Although he didn't know the exact value of $\sqrt{\pi}$, he approximated the integral well enough to predict how much money their lotteries would make.

Example: $\int_1^\infty \frac{dx}{x}$

Our next examples of indefinite integrals come close to the dividing line between infinite integral values and finite ones. We're exploring this boundary because we often want to know how large a variable value we must account for before we can ignore the rest.

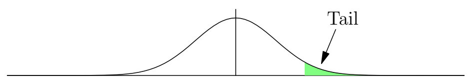


Figure 1: The tail of e^{-x^2}.

For example, when calculating probabilities the area of the "tail" of the normal curve gives the probability of an extreme value of x arising. If the area of the tail is negligible you don't have to compute it, but if it's not negligible (if it's a "fat tail") and you don't take it into account you can get a nasty surprise — like the recent mortgage scandal. The job of a mathematician is to know what finite regions require careful calculation and what areas are too small to affect the final outcome.

How "fat" does a tail have to be before its area becomes infinite and overwhelms the area of the central body? The borderline cases are x^p, where p is negative. Let's start by looking at the case for which $p = 1$; in other words:

$$\int_1^\infty \frac{dx}{x}.$$

As usual we start by computing the integral from 1 to N and then let N go to infinity.

$$\begin{aligned} \int_1^N \frac{dx}{x} &= \ln x\big|_1^N \\ &= \ln N - \ln 1 \\ &= \ln N - 0 \end{aligned}$$

As N goes to infinity so does $\int_1^N \frac{dx}{x}$, so we conclude that:

$$\int_1^\infty \frac{dx}{x} \quad \text{diverges.}$$

Example: $\displaystyle\int_1^\infty \frac{dx}{x^p}$

We know that $\int_1^\infty \frac{dx}{x}$ diverges. Next we'll find $\int_1^\infty \frac{dx}{x^p}$ for any value of p; we'll see that $p = 1$ is a borderline when we do this calculation.

$$\begin{aligned}
\int_1^\infty \frac{dx}{x^p} &= \int_1^\infty x^{-p}dx \\
&= \left.\frac{x^{-p+1}}{-p+1}\right|_1^\infty \\
&= \frac{\infty^{-p+1}}{-p+1} - \frac{1^{-p+1}}{-p+1} \\
&= \frac{\infty^{-p+1}}{-p+1} + \frac{1}{p-1}
\end{aligned}$$

Remember that the ∞ in this expression is shorthand for "a number approaching infinity".

When we think about raising a very large number to the $p+1$ power we see that there are two cases that split exactly at $p = 1$. When $p = 1$, the exponent is zero and so is the denominator; the expression doesn't make any sense. For all other values of p the expression makes sense and the value of the integral depends on whether $-p+1$ is positive or negative.

$$\frac{\infty^{-p+1}}{-p+1} \quad \text{is infinite when } -p+1 > 0$$

and

$$\frac{\infty^{-p+1}}{-p+1} \quad \text{is zero when } -p+1 < 0.$$

Check this yourself — this is the sort of problem that will be on the exam.

Conclusion: Combining this with our previous example we see that:

$$\int_1^\infty \frac{dx}{x^p} \quad \text{diverges if } p \leq 1$$

and

$$\int_1^\infty \frac{dx}{x^p} \quad \text{converges to } \tfrac{1}{p-1} \text{ if } p > 1.$$

Notice that when $p = 1$ our formula for the antiderivative is wrong; the antiderivative is $\ln x$ and not $\frac{x^{-p+1}}{-p+1}$. We really needed to do three separate calculations to compute the value of this integral: one for $p < 1$, one for $p = 1$ and one for $p > 1$.

Introduction to Limit Comparison

Recall that we're trying to find out whether the tail of a function is fat or thin — whether we can safely ignore the values of the function after a certain point or whether the area under the graph of the function is infinite. If you can't directly compute the area of the tail you can use *limit comparison* to compare your function to some function whose values you are able to compute.

If $f(x) \sim g(x)$ as x goes to infinity then $\int_a^\infty f(x)\,dx$ and $\int_a^\infty g(x)\,dx$ either both converge or both diverge.

(Remember that "$f \sim g$ as x goes to infinity" means that $\frac{f(x)}{g(x)} \to 1$ as $x \to \infty$.)

Example: $\displaystyle\int_0^\infty \frac{dx}{\sqrt{x^2+10}}$

Since $\sqrt{x^2+10} \sim \sqrt{x^2} = x$, we can compare this integral to $\int_1^\infty \frac{dx}{x}$. (Because $1/x$ is singular at $x = 0$ we'll start by ignoring the finite value $\int_0^1 \frac{dx}{\sqrt{x^2+10}}$ — if it turns out that the integral converges we can add it later.)

$$\int_1^\infty \frac{dx}{\sqrt{x^2+10}} \sim \int_1^\infty \frac{dx}{x}$$

We know that $\int_1^\infty \frac{dx}{x}$ diverges, so limit comparison tells us that $\int_0^\infty \frac{dx}{\sqrt{x^2+10}}$ diverges as well.

Question: Why did we switch from 0 to 1?

Answer: Because $\int_0^1 \frac{dx}{x}$ is infinite for unrelated reasons, and we didn't want that to affect our results. The interval between 0 and 1 is not the part of the function that we care about. What we're really worried about is what the area of the tail of the function is. We could just as well have done the comparison:

$$\int_{100}^\infty \frac{dx}{\sqrt{x^2+10}} \sim \int_{100}^\infty \frac{dx}{x}$$

which leads us to the same conclusion — the area of the tail of the graph of $\frac{1}{\sqrt{x^2+10}}$ is infinite.

Example: $\displaystyle\int_{10}^{\infty} \frac{dx}{\sqrt{x^3+3}}$

We could have used a trig substitution to compute $\int_0^\infty \frac{dx}{\sqrt{x^2+10}}$ in the previous example. We can use the limit comparison method to determine whether an integral is finite even if we're unable to find an antiderivative.

For instance, we can't evaluate $\int_{10}^\infty \frac{dx}{\sqrt{x^3+3}}$. But because:

$$\frac{1}{\sqrt{x^3+3}} \cong \frac{1}{\sqrt{x^3}} = \frac{1}{x^{3/2}}$$

we know that:

$$\int_{10}^{\infty} \frac{dx}{\sqrt{x^3+3}} \cong \int_{10}^{\infty} \frac{dx}{x^{3/2}}$$

and so we know that the integral converges to some finite value.

Example: $\int_{-\infty}^{\infty} e^{-x^2}\,dx$

We've been told that $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$. We can't compute the exact value of this integral, but *can* use a simple comparison to check that the value is finite.

We start by using the fact that this is an even function, symmetric about the y-axis, to rewrite the integral as:

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \int_{0}^{\infty} e^{-x^2}\,dx.$$

The function $f(x) = e^{-x^2}$ goes to zero so quickly that we can't find a function $g(x)$ that's comparable to $f(x)$ for a limit comparison, so we'll have to use an ordinary comparison to determine whether this improper integral converges.

Because $x^2 \geq x$ when $x \geq 1$, we know that $-x^2 \leq -x$ and $e^{-x^2} \leq e^{-x}$ for $x \geq 1$. To show that $\int_{-\infty}^{\infty} e^{-x^2}\,dx$ converges we split the integral again between $x > 1$ and $x < 1$. We compare integrals using our understanding that increasing the integrand increases the value of the integral:

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{-x^2}\,dx &= 2\int_{0}^{\infty} e^{-x^2}\,dx \\
&= 2\int_{0}^{1} e^{-x^2}\,dx + 2\int_{1}^{\infty} e^{-x^2}\,dx \\
&\leq 2\int_{0}^{1} e^{-x^2}\,dx + 2\int_{1}^{\infty} e^{-x}\,dx \quad \text{(larger integrand)}
\end{aligned}$$

Since $\int_0^1 e^{-x^2}\,dx$ is finite and $\int_1^{\infty} e^{-x}\,dx$ converges, we conclude that $\int_0^{\infty} e^{-x^2}\,dx$ converges.

Ordinary comparison is a good tool for proving the convergence of integrals whose integrands decay very rapidly.

Indefinite Integrals over Singularities

When computing $\int_0^\infty \frac{dx}{\sqrt{x^2+10}}$ we had to take an extra step to avoid the integral $\int_0^1 \frac{dx}{x}$. We'll now go back and discuss integration near singular points.

Integrals like $\int_0^1 \frac{dx}{x}$ are known as *indefinite integrals of the second type.* Examples include:

$$\int_0^1 \frac{dx}{\sqrt{x}}, \quad \int_0^1 \frac{dx}{x}, \quad \text{and} \quad \int_0^1 \frac{dx}{x^2}.$$

These integrals turn out to be fairly straightforward to calculate:

$$\begin{aligned}
\int_0^1 \frac{dx}{\sqrt{x}} &= \int_0^1 x^{-1/2}dx \\
&= \frac{1}{1/2}x^{1/2}\Big|_0^1 \\
&= 2x^{1/2}\Big|_0^1 \\
&= 2\cdot 1^{1/2} - 2\cdot 0^{1/2} \\
&= 2.
\end{aligned}$$

$$\begin{aligned}
\int_0^1 \frac{dx}{x} &= \ln x|_0^1 \\
&= \ln 1 - \ln 0 \quad \text{(diverges.)}
\end{aligned}$$

$$\begin{aligned}
\int_0^1 \frac{dx}{x^2} &= -x^{-1}\big|_0^1 \\
&= -\frac{1}{1} - \left(-\frac{1}{0}\right) \quad \text{(diverges.)}
\end{aligned}$$

However, you can get into trouble if you're not careful. Consider the following calculation:

$$\begin{aligned}
\int_{-1}^1 \frac{dx}{x^2} &= -x^{-1}\big|_{-1}^1 \\
&= -(1^{-1}) - (-(-1)^{-1}) \\
&= -1 - 1 \\
&= -2.
\end{aligned}$$

This is ridiculous! As we see from Figure 1, $\frac{1}{x^2}$ is always positive. The area under the graph of $y = \frac{1}{x^2}$ between -1 and 1 is clearly greater than 2; in particular it cannot be a negative number.

In fact, the area under the graph of $y = \frac{1}{x^2}$ between -1 and 1 is infinite, not -2. The calculation above is nonsense.

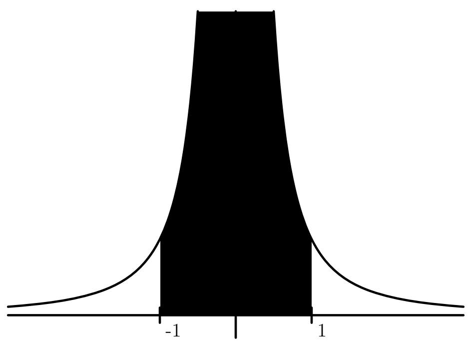


Figure 1: Graph of $y = \frac{1}{x^2}$.

Improper Integrals of the Second Kind, Continued

We'll continue our discussion of integrals of functions which have singularities at finite values; for example, $f(x) = \frac{1}{x}$. If $f(x)$ has a singularity at 0 we define

$$\int_0^1 f(x)\,dx = \lim_{a\to 0^+} \int_a^1 f(x)\,dx.$$

As before, we say the integral *converges* if this limit exists and *diverges* if not.

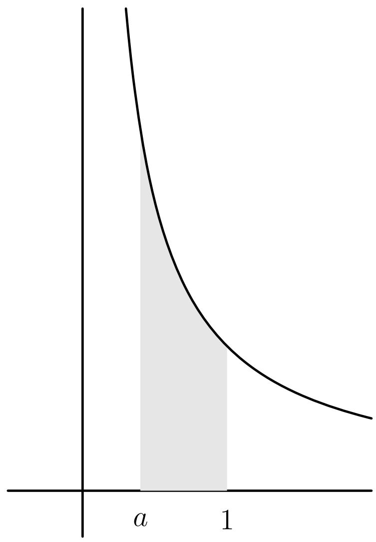


Figure 1: Area under the graph of $y = \frac{1}{x}$.

We treat this infinite vertical "tail" the same way we treated horizontal tails. Figure 1 shows a function whose value goes to positive infinity as x goes to zero from the right hand side. We don't know whether the area under its graph between 0 and 1 is going to be infinite or finite, so we cut it off at some point a where we know it will be finite. Then we let a go to zero from above ($a \to 0^+$) and see whether the area under the curve between a and 1 goes to infinity or to some finite limit.

Example: $\displaystyle\int_0^1 \frac{dx}{\sqrt{x}}$

$$\begin{aligned}\int_0^1 \frac{dx}{\sqrt{x}} &= \int_0^1 x^{-1/2}\,dx \\ &= \frac{1}{1/2} x^{1/2}\Big|_0^1 \\ &= 2x^{1/2}\Big|_0^1\end{aligned}$$

$$\begin{aligned} &= 2 \cdot 1^{1/2} - 2 \cdot 0^{1/2} \\ &= 2. \end{aligned}$$

This is a convergent integral.

Example: $\displaystyle\int_0^1 \frac{dx}{x}$

$$\begin{aligned} \int_0^1 \frac{dx}{x} &= \ln x|_0^1 \\ &= \ln 1 - \ln 0^+ \\ &= 0 - (-\infty) \\ &= +\infty. \end{aligned}$$

This integral diverges.

In general:

$$\begin{aligned} \int_0^1 \frac{dx}{x^p} &= \left.\frac{x^{-p+1}}{-p+1}\right|_0^1 \quad \text{(for } p \neq 1\text{)} \\ &= \frac{1^{-p+1}}{-p+1} - \frac{0^{-p+1}}{-p+1} \\ &= \begin{cases} \dfrac{1}{1-p} & \text{if } p < 1 \\ \text{diverges} & \text{if } p \geq 1. \end{cases} \end{aligned}$$

Overview of Improper Integrals

Now let's contrast the two types of improper integrals we've looked at — one in which x goes to infinity and one in which x approaches a point of singularity.

We have just considered functions like:

$$\frac{1}{x^{1/2}} << \frac{1}{x} << \frac{1}{x^2} \quad \text{as } x \to 0^+.$$

Conversely,

$$\frac{1}{x^{1/2}} >> \frac{1}{x} >> \frac{1}{x^2} \quad \text{as } x \to \infty.$$

In general, we found that improper integrals of functions smaller than $\frac{1}{x}$ converge while improper integrals of functions larger than or equal to $\frac{1}{x}$ diverge. Whether a function is smaller or larger than $\frac{1}{x}$ depends on the function and on what limit we're taking:

$$\frac{1}{x^{1/2}} << \underbrace{\frac{1}{x} << \frac{1}{x^2}}_{\text{integral diverges}} \quad \text{as } x \to 0^+.$$

$$\underbrace{\frac{1}{x^{1/2}} >> \frac{1}{x}}_{\text{integral diverges}} >> \frac{1}{x^2} \quad \text{as } x \to \infty.$$

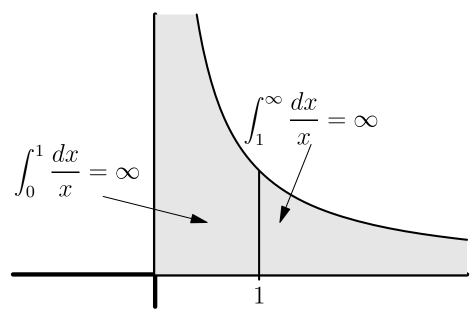


Figure 1: Area under the graph of $y = \frac{1}{x}$.

As shown in Figure 1, the graph of $f(x) = \frac{1}{x}$ is symmetric to itself by a reflection across the line $y = x$. The total area under the curve to the right of $x = 1$ is infinite and so is the area under the curve between $x = 0$ and $x = 1$.

The graph of $y = \frac{1}{x^{1/2}}$ lies below that of $y = \frac{1}{x}$ on the left and above $\frac{1}{x}$ on the right. (See Figure 3.) The total area under the graph of $y = \frac{1}{x^{1/2}}$ to the right of $x = 1$ is infinite, but the area under the curve between $x = 0$ and $x = 1$ is 2. (See Figure 2.)

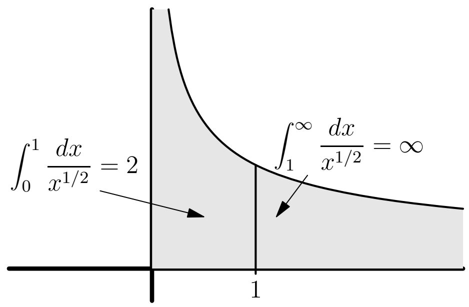


Figure 2: Area under the graph of $y = \frac{1}{x^{1/2}}$.

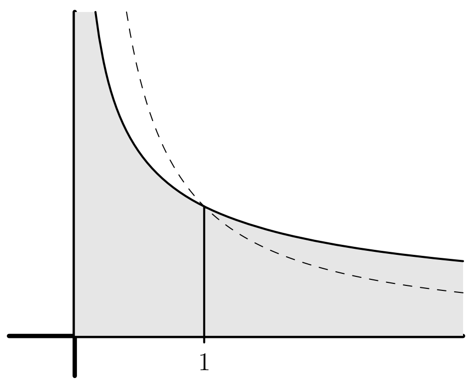


Figure 3: Graph of $y = \frac{1}{x}$ superimposed on graph of $y = \frac{1}{x^{1/2}}$.

Compare this to the area under the graph of $y = \frac{1}{x^2}$. Here the area to the right of 1 is finite (2) and the area between 0 and 1 is infinite. (See Figure 4.)

By comparing the sizes of the vertical and horizontal "tails" of the functions we can get a geometric sense of the difference between convergent and divergent indefinite integrals.

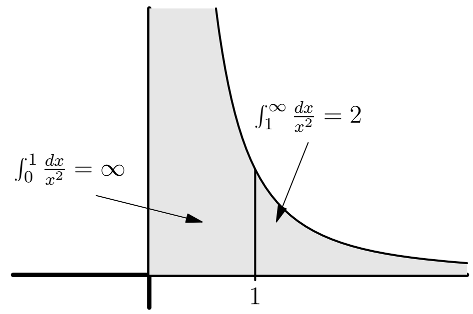


Figure 4: Area under the graph of $y = \frac{1}{x^2}$.

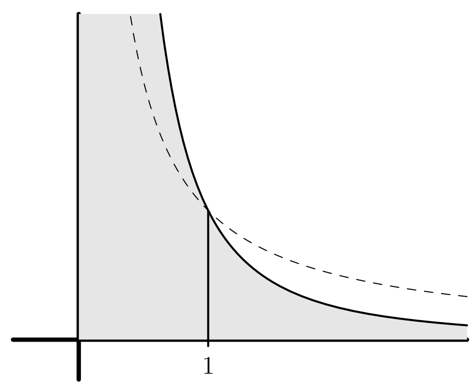


Figure 5: Graph of $y = \frac{1}{x}$ superimposed on graph of $y = \frac{1}{x^2}$.

An Improper Integral of the Second Kind

Suppose we want to calculate:

$$\int_0^\infty \frac{dx}{(x-3)^2}.$$

In calculating $\int_0^\infty \frac{dx}{(x-3)^2}$ you must worry about two pieces — say $\int_0^5 \frac{dx}{(x-3)}^2$ and $\int_5^\infty \frac{dx}{(x-3)}^2$.

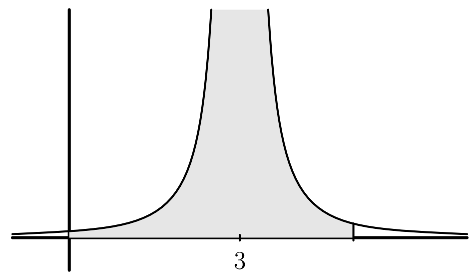


Figure 1: $\int_0^5 \frac{dx}{(x-3)^2} = \infty$.

The singularity in the graph of $y = \frac{dx}{(x-3)^2}$ is comparable to that of $y = \frac{1}{x^2}$ near $x = 0$. The area under the graph of $y = \frac{dx}{(x-3)^2}$ between 0 and 5 is infinite.

However, $\int_5^\infty \frac{dx}{(x-3)}^2$ converges.

Unfortunately, the total $\int_0^\infty \frac{dx}{(x-3)^2}$ diverges because it's the sum of a divergent indefinite integral and a convergent one.

If you failed to notice the singularity at $x = 3$ you might have calculated the value of this integral to be finite, the same way we calculated the false value of $\int_{-1}^1 \frac{dx}{x^2}$. In that case you might still be saved from a terrible error by noticing that you calculated a negative value for the integral of a function that is everywhere positive.

Question: How do we know when and where to look for "bad spots" in an integrand?

Answer: If the either limit of integration is infinite, check the limits as x goes to infinity and/or minus infinity. Also, check any singularity, like x going to 3 in the integrand $\frac{1}{(x-3)^2}$. In other words, look in all the places where the integrand may be infinite. Once you've identified the problem areas, you can focus on each one separately by splitting the domain of integration into parts.

Specific things to watch for are negative exponents — anything of the form $\frac{dx}{x^n}$ goes to infinity as x approaches zero. This includes all the integrals we computed using partial fractions; whenever there's something in the denominator that can be zero, there's a singularity.

Introduction to Series

Remember that we're dealing with infinity; we'll be studying *infinite* series. The most important and useful series is the geometric series. We'll start with a concrete example:

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \cdots$$

We can visualize this infinite sum by marking its partial sums on the number line, as shown in Figure 1. The figure shows the results of adding 1, then $\frac{1}{2}$, then $\frac{1}{4}$, then $\frac{1}{8}$. Note that the value of each partial sum is midway between the value of the previous sum and 2. We say that the series converges to 2.

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \cdots = 2.$$

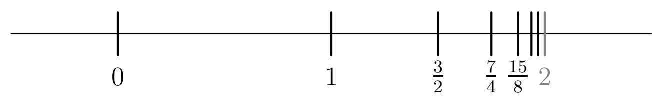


Figure 1: Adding a term gives a number half way between the previous value and 2.

Note that we never get to 2, we just get halfway there infinitely often. This is known as Zeno's paradox — if a rabbit gives a tortoise a head start in a race, the rabbit can never pass the tortoise because it must cross half the distance between itself and the tortoise infinitely many times. One resolution of this paradox requires understanding time as a continuum, which Zeno failed to do.

In general, a *geometric series* is a series of the form:

$$1 + a + a^2 + a^3 + \cdots = \frac{1}{1-a} \quad (|a| < 1).$$

We'll see how this formula is derived later.

Divergent Series

As with indefinite integrals, we're concerned about when infinite series converge. We're also interested in what goes wrong when a series diverges — when it fails to converge. Recall that when $|a| < 1$,

$$1 + a + a^2 + a^3 + \cdots = \frac{1}{1-a}.$$

How does this fail when $|a| \geq 1$?

One simple example of a divergent series is a geometric series with a equal to 1:

$$1 + 1 + 1 + 1 + \cdots = \frac{1}{1-1} = \frac{1}{0}.$$

This almost makes sense! Since the sum is infinite we conclude that the series diverge.

Now let's try $a = -1$. We get:

$$1 - 1 + 1 - 1 + \cdots$$

If we look at the partial sums of this sequence we see that they alternate between 1 and 0. If we plug $a = -1$ into our formula $\frac{1}{1-a}$ it predicts that the sum is $\frac{1}{2}$ which is halfway between 0 and 1 but is still wrong. Using the formula here is the equivalent of computing an indefinite integral without checking for singularities; it gives you an interesting but wrong result.

Because the partial sums of the series alternate between 0 and 1 without ever tending toward a single number, we say that this series is also divergent. The geometric series only converges when $|a| < 1$.

We'll look at one more case: $a = 2$. According to the formula,

$$1 + 2 + 2^2 + 2^3 + \cdots = \frac{1}{1-2} = -1.$$

This is clearly wrong. The sequence diverges; the left hand side is obviously infinite and the right hand side is -1. But number theorists actually have a way of making sense out of this, if they're willing to give up on the idea that 0 is less than 1.

Notation for Series

It's easier to understand and explain mathematics if we have good notation for what we're discussing. In this case, we define a *partial sum* to be:

$$S_N = \sum_{n=0}^{N} a_n.$$

Now we can do something similar to what we did for indefinite integrals and define:

$$S = \sum_{n=0}^{\infty} a_n = \lim_{N \to \infty} S_N.$$

Once again we have two choices. If the limit exists we say that the series converges. If the limit does not exist we say that the series diverges.

Examples of Series

Example: $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$

We'll now look at some other interesting series and will return to the (very important) geometric series later.

The series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ turns out to be very similar to the improper integral $\int_1^{\infty} \frac{dx}{x^2}$, which is convergent. The series is also convergent.

It's easy to calculate that $\int_1^{\infty} \frac{dx}{x^2} = 1$. It's hard to calculate that:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

This was computed by Euler in the early 1700's.

Example: $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^3}$

Notice that the series in these examples start with $n = 1$; we can't start with $n = 0$ because a_0 would be of the form $\frac{1}{0}$. That's ok — we can start these series at any positive value of n and it won't make a difference to whether they converge or diverge.

The series $\sum_{n=1}^{\infty} \frac{1}{n^3}$ is similar to the integral $\int_1^{\infty} \frac{dx}{x^3}$, which converges to $\frac{1}{2}$. The series $\sum_{n=1}^{\infty} \frac{1}{n^3}$ was only recently proved to converge to an irrational number; there is no elementary way of describing its value.

We've been loosely comparing series to integrals; we can make this formal by using a Riemann sum with $\Delta x = 1$.

Comparison of the Harmonic Series

We're going to compare:

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

to $\displaystyle\int_1^{\infty} \frac{dx}{x}$, using Riemann sums to show that the series diverges. The same sort of reasoning is applicable to the previous two examples.

We'll use a Riemann sum to calculate the area under the graph of $y = \frac{1}{x}$ using $\Delta x = 1$. We also get to choose whether to compute the upper Riemann sum or lower; we'll do both.

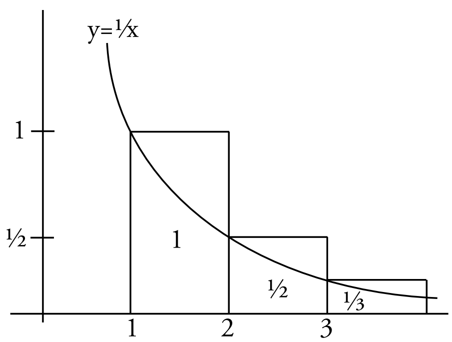


Figure 1: Upper Riemann sum; $y = \frac{1}{x}$ (not to scale).

The upper Riemann sum is the sum of the areas of the rectangles indicated in Figure 1. If we stop at the rectangle just before N we see that the area under the curve from 1 to N is $\int_0^N \frac{dx}{x}$ and that that's less than the sum of the areas of the rectangles:

$$\int_1^N \frac{dx}{x} < 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N-1}.$$

(There are only $N-1$ rectangles here because the distance between 1 and N is $N-1$.) If $S_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N}$, then because there's one more term in the sum we have:

$$\int_1^N \frac{dx}{x} < 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N-1} < S_N.$$

Thus we know that the value of the integral is less than the value of S_N. This will allow us to prove conclusively that the series diverges.

$$\int_1^N \frac{dx}{x} \quad < \quad S_N$$

$$\begin{aligned} \ln x|_1^N &< S_N \\ \ln N - \ln 1 &< S_N \\ \ln N - 0 &< S_N \\ \ln N &< S_N \end{aligned}$$

As N goes to infinity, $\ln N$ goes to infinity and so S_N must also go to infinity. By definition,

$$\sum_{n=1}^{\infty} \frac{1}{n} = \lim_{N\to\infty} S_N,$$

so the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.

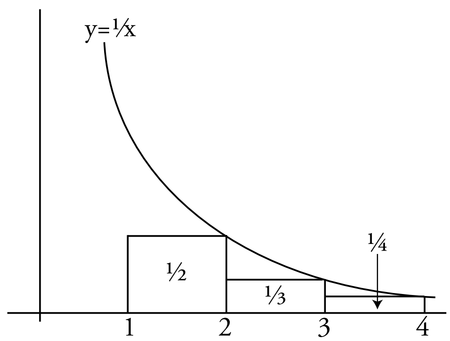


Figure 2: Lower Riemann sum; $y = \frac{1}{x}$ (not to scale).

Now we'll use the lower Riemann sum to see that S_N goes to infinity at the same rate as $\int_1^N \frac{dx}{x}$. Since the sum of the areas of the rectangles shown in Figure 2 is less than the area under the curve, we know that:

$$\int_1^N \frac{dx}{x} > \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N} = S_N - 1.$$

Recall that $\int_1^N \frac{dx}{x} = \ln N$, so we get:

$$\begin{aligned} S_N - 1 &< \int_1^N \frac{dx}{x} \\ S_N &< (\ln N) + 1. \end{aligned}$$

Combining this with the result from the upper Riemann sum, we conclude:

$$\ln N < S_N < (\ln N) + 1.$$

The value of S_N is hard to calculate exactly, but now we know that it's between $\ln N$ and $(\ln N) + 1$, which we can compute.

Comparison Tests

Integral Comparison

We used integral comparison when we applied Riemann sums to understanding $\sum_1^\infty \frac{1}{n}$ in terms of $\int_1^\infty \frac{dx}{x}$, and we've made several other comparisons between integrals and series in this lecture. Now we learn the general theory behind this technique.

Theorem: If $f(x)$ is decreasing and $f(x) > 0$ on the interval from 1 to infinity, then either the sum $\sum_1^\infty f(n)$ and the integral $\int_1^\infty f(x)\,dx$ both diverge or they both converge and:

$$\sum_{n=1}^{\infty} f(n) - \int_1^\infty f(x)\,dx \quad < \quad f(1).$$

For example, when $S_N = \sum_1^N \frac{1}{n}$ we showed that $|S_n - \ln N| < 1$.

Since it's very difficult to compute infinite sums and it's easy to compute indefinite integrals, this is an extremely useful theorem.

Limit Comparison

Theorem: If $f(n) \sim g(n)$ (i.e. if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$) and $g(n) > 0$ for all n, then either both $\displaystyle\sum_{n=1}^{\infty} f(n)$ and $\displaystyle\sum_{n=1}^{\infty} g(n)$ converge or both diverge.

This says that if f and g behave the same way in their tails, their convergence properties will be similar.

Examples of Comparison

Example: $\displaystyle\sum_{n=0}^{\infty} \frac{1}{\sqrt{n^2+1}}$

We know that $\frac{1}{\sqrt{n^2+1}}$ is comparable to $\frac{1}{\sqrt{n^2}} = \frac{1}{n}$, so by limit comparison we know that $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^2+1}}$ converges or diverges as $\sum_{n=0}^{\infty} \frac{1}{n}$ does. We proved earlier that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges, so $\sum_{n=0}^{\infty} \frac{1}{\sqrt{n^2+1}}$ must diverge as well.

Note that we can include the $n = 0$ term in $\sum_{n=0}^{\infty} \frac{1}{\sqrt{n^2+1}}$ but not in $\sum_{n=1}^{\infty} \frac{1}{n}$. This is ok; the limit comparison test is is concerned only with long term behavior, not with the early partial sums.

Example: $\displaystyle\sum_{n=2}^{\infty} \frac{1}{\sqrt{n^5-n^2}}$

Because we have a subtraction in the denominator we have to be careful not to divide by zero; we start our series at $n = 2$.

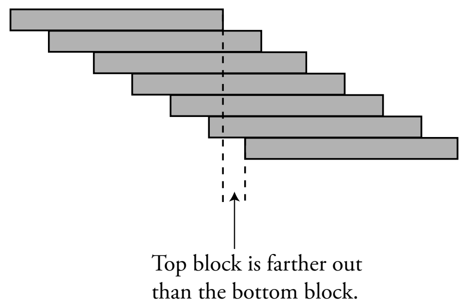
We compare $\frac{1}{\sqrt{n^5-n^2}}$ to $\frac{1}{\sqrt{n^5}} = \frac{1}{n^{5/2}}$. If we choose the right function to be the numerator, it's relatively simple to show that they are similar:

$$\begin{aligned}
\lim_{n\to\infty} \frac{f(n)}{g(n)} &= \lim_{n\to\infty} \frac{\frac{1}{\sqrt{n^5}}}{\frac{1}{\sqrt{n^5-n^2}}} \\
&= \lim_{n\to\infty} \frac{\sqrt{n^5-n^2}}{\sqrt{n^5}} \\
&= \lim_{n\to\infty} \sqrt{\frac{n^5}{n^5} - \frac{n^2}{n^5}} \\
&= \sqrt{1-0} \\
&= 1.
\end{aligned}$$

Since the two functions are similar, we can apply the limit comparison test and conclude that because $\sum_{n=1}^{\infty} \frac{1}{n^{5/2}}$ converges ($\frac{5}{2} > 1$), $\sum_{n=2}^{\infty} \frac{1}{\sqrt{n^5-n^2}}$ must also converge.

Preview of Stacking Blocks

Question: Is it possible to stack Professor Jerison's blocks so that no part of the top block is above the bottom block?



He answers this question in the next lecture.

Stacking Blocks

Is it possible to stack blocks as shown in Figure 1, so that no part of the bottom block is below the top block? In general, how much horizontal distance can there be between the top block and the bottom block?

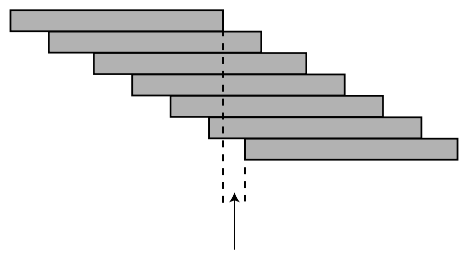


Figure 1: Stack of blocks.

This is a good kind of math question to ask. If there's a limit to how distant the top block can be from the bottom block it will be interesting to know what it is. It would also be interesting to discover that there's no limit to this distance. In the end, this becomes a question of whether the top block's position converges or diverges.

Professor Jerison has eight blocks; do you think he can achieve his goal with what he has?

In order to get the greatest possible horizontal distance, start at the top of the stack and work downward. The topmost block has to cover at least half of the block below it or else it will fall. Slide it so that its right end is at the midpoint of the block below it.

Next, slide the second block down as far to the left as you can without upsetting the tower. Then slide the block below it to the left as far as you can, and the one below that, and the one below that, and so on. In this way, Professor Jerison accomplished his goal using only 7 blocks.

How much further could we get if we had more blocks? Let's calculate it.

Calculation

To make the calculations simple, let's say that each block is 2 units long. Then if the left end of the topmost block is at position 0, then the left end of the block under it is at position 1.

In general the center of mass of the top n blocks must always be above the block supporting them.

- Let C_1 = the x coordinate of the center of mass of the top block.

- Let $C_2 =$ the x coordinate of the center of mass of the top two blocks.
- Put the left end of the next block below the center of mass of the previous ones.

Question: How do you know that this is the best way to stack them?

Answer: I can't answer that in general, but I can tell you that this is the best we can do if we start building from the top – we're using what computer scientists call "the greedy algorithm" and going as far as we can at each step.

If we tried using this algorithm starting from the bottom it wouldn't work. We'd stack the second block with its end on the midpoint of the first block and then be unable to gain any distance beyond that.

It seems possible that there is some other strategy that's better than using the greedy algorithm starting at the top. There isn't, but we're not going to prove that today.

According to our strategy we need to know C_N, the x coordinate of the center of mass of the top N blocks, in order to continue.

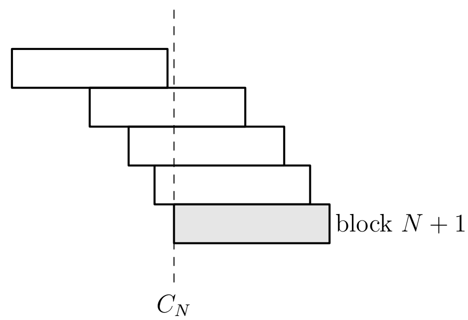


Figure 2: Adding a block.

If the center of mass of the top N blocks is on the line $x = C_N$, the center of mass of the $(N+1)^{st}$ block will have x coordinate $C_N + 1$. This shifts the center of mass of the stack to the right; the x coordinate of the new center of mass of the top $N+1$ blocks is given by the weighted average of the centers of mass of the stack:

$$\begin{aligned} C_{N+1} &= \frac{NC_N + 1(C_N+1)}{N+1} \\ &= \frac{(N+1)C_N + 1}{N+1} \\ C_{N+1} &= C_N + \frac{1}{N+1}. \end{aligned}$$

Adding the $(N+1)^{st}$ block added to the stack allows you to extend the stack $\frac{1}{N+1}$ units farther from its base.

So

$$\begin{aligned} C_1 &= 1 \\ C_2 &= 1 + \frac{1}{2} \\ C_3 &= 1 + \frac{1}{2} + \frac{1}{3} \\ C_4 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \\ C_5 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} > 2. \end{aligned}$$

It takes at least 5 blocks to extend the top block beyond the base.

$$C_N = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{N}$$

This sum C_N is the same as S_N from a previous lecture:

$$C_N = S_N = \sum_{n=1}^{N} \frac{1}{n}.$$

We know that:

$$\ln N < S_N < (\ln N) + 1.$$

Since $\ln N$ goes to infinity as N goes to infinity, $S_N = C_N$ must go to infinity as N does. If we have enough blocks we *can* extend our stack as far as we want.

In this example, the fact that $\sum_{n=1}^{N} \frac{1}{n}$ diverges means that it's possible to extend the stack as far to the left as we wish, provided we have enough blocks.

On the other hand, the inequality $S_N < (\ln N) + 1$ tells us that it will take a lot of blocks to extend the top of the stack very far.

How high would this stack of blocks be if it extended across the two lab tables at the front of the lecture hall? One lab table is 6.5 blocks, or 13 units, long. Two tables are 26 units long. There will be $26 - 2 = 24$ units of overhang in the stack. (We subtract 2 because the bottom block has no overhang and because the stack extends one unit past the center of mass of the top block.) Each block is approximately 3 centimeters tall.

If $\ln n = 24$ then $n = e^{24}$ and:

$$\text{Height} = 3cm \cdot e^{24} \approx 8 \times 10^8 m.$$

That height is roughly twice the distance to the moon.

If you want the stack to span this room (~ 30 ft.) it would have to be 10^{26} meters high. That's about the diameter of the observable universe.

We can learn one more thing from this experiment — if we look at the stack sideways we see that it follows the shape of the graph of $\ln x$. This experiment provides a concrete example of how slowly the function $\ln x$ increases.

We did not discover an important number that limited the reach of the stack, but we did discover that that reach is infinite — infinity is also an important number. We also discovered a property of that infinite value; that the rate of extension of the stack is very slow. Infinity doesn't have a single value; there are lots of different orders of infinity.

Power Series

Our last subject will be *power series*. We've seen one power series:

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x} \quad (|x| < 1).$$

This is our geometric series, with x in place of a. We'll now see why the sum should equal $\frac{1}{1-x}$.

Suppose that:

$$1 + x + x^2 + x^3 + \cdots = S$$

for some number S. Multiply both sides of this equation by x:

$$x + x^2 + x^3 + x^4 + \cdots = Sx.$$

Now subtract the two equations.

$$\begin{array}{ccccccccccc} 1 & + & x & + & x^2 & + & x^3 & + & \cdots & = & S \\ & & x & + & x^2 & + & x^3 & + & \cdots & = & Sx \\ \hline 1 & + & 0 & + & 0 & + & 0 & + & \cdots & = & S - Sx \end{array}$$

Lots of terms cancel! Continuing, we get:

$$\begin{aligned} 1 &= S - Sx \\ 1 &= S(1-x) \\ \frac{1}{1-x} &= S. \end{aligned}$$

There is a flaw in this reasoning — the argument only works if S exists. For example, if $x = 1$ this technique tells us that $\infty - \infty = \infty - \infty$. This is not a useful result.

This line of reasoning leads to a correct answer exactly when the series converges; in other words, when $|x| < 1$.

General Power Series

We now know exactly which geometric series converge and, if they do converge, what they converge to. Our next step is to extend this result to cover all power series. In general, a *power series* is a series of the form:

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots = \sum_{n=0}^{\infty} a_n x^n.$$

As with geometric series, there's a very simple rule that tells us when power series converge. The series converges when $-R < x < R$ for some "magic number" R called the *radius of convergence*. The value of R depends on the values of the coefficients a_i. When $|x| > R$, the sum $\sum a_n x^n$ diverges. When $|x| = R$ the series might or might not converge — we won't deal with this case in this course.

What always happens is that for $|x| < R$, the values $|a_n x^n|$ tend to 0 exponentially fast. When $|x| > R$, the values $|a_n x^n|$ won't tend to 0 at all. For example, when $x = 1/2$ the geometric series looks like:

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$$

We're adding together numbers that get smaller and smaller. When $x = 2$ the numbers in our summation get bigger and bigger:

$$1 + 2 + 4 + 8 + \cdots$$

Question: How can you tell when the numbers $a_n x^n$ are declining *exponentially* fast?

Answer: Any time the power series converges the numbers $a_n x^n$ decline exponentially fast. This is because of the power x^n in each term.

Question: How do you find R?

Answer: There's a long discussion of this question in many textbooks, but you won't need it. The radius of convergence will always be 1 or infinite or as obvious as it is for the geometric series.

Introduction to Taylor Series

Why are we looking at power series? If we reverse the equation for the geometric series:

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}$$

we get a description of $\frac{1}{1-x}$ in terms of a series. In fact, we can represent all of the functions we've encountered in this course in terms of series.

The technique is similar to the use of a decimal expansion to represent 1/3 or $\sqrt{2}$. When we describe a function like e^x or $\arctan x$ in terms of a series we can approximate and manipulate those functions as easily as we do polynomial functions.

Rules for convergent power series

What sorts of manipulations might we want to perform? Addition, multiplication, division, substitution (composition), integration and differentiation. The rules for manipulation of power series are essentially the same as those for manipulating polynomials!

$$f(x) + g(x), \quad f(x) \cdot g(x), \quad f(g(x)), \quad f(x)/g(x), \quad \frac{d}{dx}f(x), \quad \int f(x)dx$$

We can do all of these with power series; in this class integration and differentiation will be the most interesting manipulations.

We take the derivative of a power series just as we do for polynomials:

$$\frac{d}{dx}(a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots) = a_1 + 2a_2x + 3a_3x^2 + \cdots$$

Similarly, the formula for the integral of a power series is:

$$\int (a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots)dx = c + a_0x + \frac{a_1x^2}{2} + \frac{a_2x^3}{3} + \frac{a_3x^4}{4} + \cdots$$

Here the arbitrary constant c takes the place of the constant term in the new series.

Question: Is that a series or a polynomial?

Answer: It's a polynomial if it ends; if it goes on infinitely far then it's a series.

Question: You can add up terms of x in a series?

Answer: When we introduced series we described them as infinite sums of *numbers*. At the start of this class we rewrote the geometric series using the *variable* x in place of the "constant value" a. When we plug in a value for x we get a sum of infinitely many numbers, so as long as we remember that x is a placeholder for a numerical value there's no problem.

In other words, what we're working with here are functions of x. These functions are defined for values of x inside the radius of convergence and undefined for values of x that are too large, just as the function $f(x) = \sqrt{x}$ is defined for positive values of x and undefined for $x < 0$.

Power series are infinite sums of powers of x, with coefficients. People also study and use series that are infinite sums of sines and cosines and lots of other series, but we're only going to study power series here.

Taylor's Formula

Taylor's formula describes how to get power series representations of functions. The function e^x doesn't look like a polynomial; we have to figure out what the values of a_i have to be in order to describe e^x as a series.

Taylor's formula says that given any function f for which the n^{th} derivative $f^{(n)}(x)$ exists for x near 0,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

We'll learn how to use it soon.

Why should this work? Suppose that:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots$$

Then:

$$\begin{aligned} f'(x) &= a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \cdots \quad \text{and} \\ f''(x) &= 2a_2 + 3 \cdot 2a_3 x + 4 \cdot 3a_4 x^2 + \cdots \quad \text{and} \\ f^{(3)}(x) &= 3 \cdot 2a_3 + 4 \cdot 3 \cdot 2a_4 x + \cdots \end{aligned}$$

Evaluating each of these at 0 we see that: $f(0) = a_0$, $f'(0) = a_1$, $f''(0) = 2a_2$ and $f^{(3)}(0) = 3 \cdot 2a_3$. Solving for a_3 we get $a_3 = \dfrac{f^{(3)}(0)}{3 \cdot 2 \cdot 1}$ and in general:

$$a_n = \frac{f^{(n)}(0)}{n!},$$

where:

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 1.$$

We define $0! = 1$ because that makes our formulas work nicely.

Taylor's Formula

Recall that Taylor's formula says:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

Example: e^x

If $f(x) = e^x$ then $f'(x) = e^x$, $f''(x) = e^x$, and so on. This means that $f^{(n)}(0) = e^0 = 1$ for any n. Taylor's formula tells us that:

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

In particular, we know that:

$$\begin{aligned} e^1 &= \sum_{n=0}^{\infty} \frac{1}{n!}, \quad \text{or:} \\ e &= 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \cdots \end{aligned}$$

We finally have a straightforward way to compute the value of e.

Example: $\sin x$, $\cos x$

You may recall from the sessions on linear and quadratic approximation that

$$\sin x \approx x \quad \text{and}$$

$$\cos x \approx 1 - \frac{x^2}{2}.$$

We can use Taylor's formula to complete these formulas.

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

You may feel that these are hard to memorize, especially since we skipped some steps in their derivation. Try to use their similarities to help you remember them. The exponential function has a power series in which all the $a_i = 1$. The power series expansion of sine has all the odd powers with alternating signs. The series for cosine has all the even powers with alternating signs. All three functions are part of the same family.

Review of Taylor's Series

Professor Jerison was away for this lecture, so Professor Haynes Miller took his place.

A *power series* or *Taylor's series* is a way of writing a function as a sum of integral powers of x:

$$f(x) = a_0 + a_x x + a_2 x^2 + \cdots$$

Polynomials are power series; they go on for a finite number of terms and then end, so that all of the a_j equal 0 after a certain point. Since polynomials are a special type of power series, it's not surprising that power series behave almost exactly like polynomials.

Given a power series $f(x) = \displaystyle\sum_{n=0}^{\infty} a_n x^n$ there is a number R $(0 \le R \le \infty)$ for which, when $|x| < R$, the sum $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ converges and when $|x| > R$ the sum diverges. R is called the *radius of convergence*.

For $|x| < R$, $f(x)$ has all its higher derivatives, and Taylor's formula tells us that $a_n = \dfrac{f^{(n)}(0)}{n!}$. So:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots$$

Whenever you write out a power series you should say what the radius of convergence is. The radius of convergence of this series is infinity; in other words, the series converges for any value of x.

Example: (Due to Leonhard Euler) e^x

We know that if $f(x) = e^x$ then $f^{(n)}(x) = e^x$ for all n, and so $f^{(n)}(0) = 1$. Applying Taylor's formula we see that:

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots, \quad R = \infty.$$

Question: How many terms of the series do we need to write out?

Answer: Write out enough terms so that you can see what the pattern is.

Question: What functions can be written as power series?

Answer: Any function that has a reasonable expression can be written as a power series. This is not a very precise answer because the true answer is a little bit complicated. For now, it's enough that any of the functions that occur in calculus (like sines cosines, and tangents) all have power series expansions.

Taylor's Series of $\dfrac{1}{1+x}$

Our next example is the Taylor's series for $\frac{1}{1+x}$; this series was first described by Isaac Newton. Remember the formula for the geometric series:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots \quad \text{if } |x| < 1.$$

If we replace x by $-x$ we get:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots \quad R = 1.$$

You may recall that the graph of this function has an infinite discontinuity at $x = -1$; this gives us an idea of what R might be. If we try to replace x by -1 we get something of the form $\infty = \infty$; the radius of convergence of this series is 1.

Instead of deriving this from the formula for the geometric series we could also have computed it using Taylor's formula. Try it!

Question: If you put in -1 for x the series diverges. If you put in 1, it looks like it would converge.

Answer: The graph of $y = \frac{1}{1+x}$ looks smooth at $x = 1$, but there is still a problem. If the series converges for $|x| < |a|$ and then diverges for $x = a$ the radius of convergence is a; that's it.

What happens if we plug $x = 1$ into the series $1 - x + x^2 - x^3 + \cdots$? Let's look at the partial sums $S_N = \sum_{n=0}^{N} a_n x^n$.

$$\begin{aligned} S_0 &= 1 \\ S_1 &= 0 \\ S_2 &= 1 \\ S_3 &= 0 \\ S_4 &= 1 \\ &\vdots \end{aligned}$$

Even though these don't go off to infinity, they still don't converge.

Taylor's Series of $\sin x$

In order to use Taylor's formula to find the power series expansion of $\sin x$ we have to compute the derivatives of $\sin(x)$:

$$\begin{aligned}
\sin'(x) &= \cos(x) \\
\sin''(x) &= -\sin(x) \\
\sin'''(x) &= -\cos(x) \\
\sin^{(4)}(x) &= \sin(x).
\end{aligned}$$

Since $\sin^{(4)}(x) = \sin(x)$, this pattern will repeat.

Next we need to evaluate the function and its derivatives at 0:

$$\begin{aligned}
\sin(0) &= 0 \\
\sin'(0) &= 1 \\
\sin''(0) &= 0 \\
\sin'''(0) &= -1 \\
\sin^{(4)}(0) &= 0.
\end{aligned}$$

Again, the pattern repeats.

Taylor's formula now tells us that:

$$\begin{aligned}
\sin(x) &= 0 + 1x + 0x^2 + \frac{-1}{3!}x^3 + 0x^4 + \cdots \\
&= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots
\end{aligned}$$

Notice that the signs alternate and the denominators get very big; factorials grow very fast.

The radius of convergence R is infinity; let's see why. The terms in this sum look like:

$$\frac{x^{2n+1}}{(2n+1)!} = \frac{x}{1} \cdot \frac{x}{2} \cdot \frac{x}{3} \cdots \frac{x}{(2n+1)}.$$

Suppose x is some fixed number. Then as n goes to infinity, the terms on the right in the product above will be very, very small numbers and there will be more and more of them as n increases.

In other words, the terms in the series will get smaller as n gets bigger; that's an indication that x may be inside the radius of convergence. But this would be true for any fixed value of x, so the radius of convergence is infinity.

Why do we care what the power series expansion of $\sin(x)$ is? If we use enough terms of the series we can get a good estimate of the value of $\sin(x)$ for any value of x.

This is very useful information about the function $\sin(x)$ but it doesn't tell the whole story. For example, it's hard to tell from the formula that $\sin(x)$ is periodic. The period of $\sin(x)$ is 2π; how is this series related to the number π?

Power series are very good for some things but can also hide some properties of functions.

Power Series Multiplication

Once you have one power series, there are ways to get new power series from it. One thing you can do is multiply — can we use power series to multiply x by $\sin(x)$?

We have a power series for $\sin(x)$:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

Can we get one for x? Yes!

$$x = 0 + x + 0x^2 + 0x^3 + 0x^4 + \cdots$$

We can treat power series just like polynomials and multiply them together:

$$\begin{aligned} x \cdot \sin(x) &= x \cdot (x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots) \\ &= x^2 - \frac{x^4}{3!} + \frac{x^6}{5!} - \frac{x^8}{7!} + \cdots \end{aligned}$$

The radius of convergence will be the smaller of the two radii of convergence; in this case $R = \infty$.

Power series multiplication is just like polynomial multiplication. It can get tedious if you're dealing with a lot of terms, but this example was pretty simple.

Notice that we just multiplied two odd functions, $\sin x$ and x, and so their product is even. That's reflected in the fact that all the terms in the power series expansion of $x \sin(x)$ have even degree. For an odd function, like $\sin(x)$, all the terms in the power series have odd degree. In general, power series of even functions will have only even degree terms and power series of odd functions will consist of all odd degree terms.

Question: Why is the radius of convergence the smaller of the two radii?

Answer: If $|x|$ is larger than the smallest radius of convergence then one of your functions isn't defined at x, so you can't expect multiplication by that function to make sense.

Derivative of a Power Series

We can differentiate power series. For example, $\cos(x) = \sin'(x)$ so we can find a power series for $\cos(x)$ by differentiating the power series for $\sin(x)$ term by term — the same way we differentiate polynomials.

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$\begin{aligned}
\cos(x) &= \sin'(x) \\
&= 1 - 3\frac{x^2}{3!} + 5\frac{x^4}{5!} - 7\frac{x^6}{7!} + \cdots \\
\cos(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots
\end{aligned}$$

Notice how $3\frac{x^2}{3!}$ became $\frac{x^2}{2!}$ when we canceled the 3's. This happens with each term of the power series.

The radius of convergence of the derivative of a power series is the same as the radius of convergence of the power series you started with. Here $R = 1$.

Of course, you could get this same formula using Taylor's formula and the derivatives of the cosine function.

Integral of a Power Series

We can multiply, add and differentiate power series. Can we integrate them? Yes; as you'd expect, integration of power series is very similar to integration of polynomials. We'll use integration to find a power series expansion for:

$$\ln(1+x) = \int_0^x \frac{dt}{1+t} \quad (x > -1).$$

We know that:

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \cdots$$

So:

$$\begin{aligned} \ln(1+x) &= \int_0^x (1 - t + t^2 - t^3 + \cdots)dt \\ &= \left[t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \cdots\right]_0^x \\ \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \end{aligned}$$

Because we began with a power series whose radius of convergence was 1, the radius of convergence of the result will also be 1. This reflects the fact that $\ln(1+x)$ is undefined for $x \leq -1$.

Question: If you only use positive values of x is there still a radius of convergence?

Answer: Yes. If $x > 1$ then the numerators x, x^2, x^3, x^4 and so on are increasing exponentially. The denominators 1, 2, 3, 4 ... only grow linearly. So as n goes to infinity, $\frac{x^n}{n}$ will also go to infinity. If the terms of a series go to infinity then the series diverges.

Euler used this kind of power series expansion to calculate natural logarithms much more efficiently than was previously possible.

Substitution of Power Series

We can find the power series of e^{-t^2} by starting with the power series for e^x and making the substitution $x = -t^2$.

$$
\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad (R = \infty) \\
e^{-t^2} &= 1 + (-t^2) + \frac{(-t^2)^2}{2!} + \frac{(-t^2)^3}{3!} + \cdots \\
&= 1 - t^2 + \frac{t^4}{2!} - \frac{t^6}{3!} + \cdots
\end{aligned}
$$

The signs of the terms alternate, the powers are all even, and the denominators are the factorials shown. The radius of convergence is infinity.

Power Series Expansion of the Error Function

Several times in this course we've seen the error function:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt.$$

(The factor of $\frac{2}{\sqrt{\pi}}$ guarantees that $\lim_{x\to\infty} \text{erf}(x) = 1$.) This function is very important in probability theory, but we don't have a conventional algebraic description of it.

Because we can integrate power series and know that:

$$e^{-t^2} = 1 - t^2 + \frac{t^4}{2!} - \frac{t^6}{3!} + \cdots \quad (R = \infty),$$

we can now find a power series expansion for the error function.

$$\begin{aligned}
\text{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt \\
&= \frac{2}{\sqrt{\pi}} \int_0^x (1 - t^2 + \frac{t^4}{2!} - \frac{t^6}{3!} + \cdots)dt \\
&= \frac{2}{\sqrt{\pi}} \left[t - \frac{t^3}{3} + \frac{t^5}{5\cdot 2!} - \frac{t^7}{7\cdot 3!} + \cdots \right]_0^x \\
&= \frac{2}{\sqrt{\pi}} \left(x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \frac{x^7}{7\cdot 3!} + \cdots \right)
\end{aligned}$$

To get this to look exactly like a power series we would distribute the factor of $\frac{2}{\sqrt{\pi}}$ across the sum, multiplying it by each term of the series. However, that's not strictly necessary.

This turns out to be a very good way to compute the value of the error function; your calculator probably uses this method.

Finale

We encourage you to go on to 18.02: Multivariable Calculus, just in case you were thinking of stopping here.

In Multivariable Calculus you'll learn about vectors and more, and you'll also see a lot of things from this course in a larger context. For example, you'll learn that our strange formulas like the product rule and quotient rule are all special cases of the chain rule. This inspired Professor Jerison to write the following poem:

> One thing to rule them all
> One thing to find them
> One thing to bring them all
> And in mathematics bind them.

(Sincere apologies to JRR Tolkien.)