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.

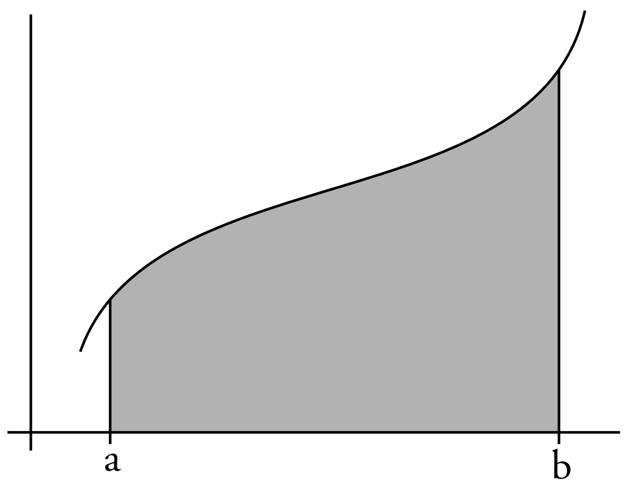

Figure 1: Area under a curve

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 $a < x < b$.

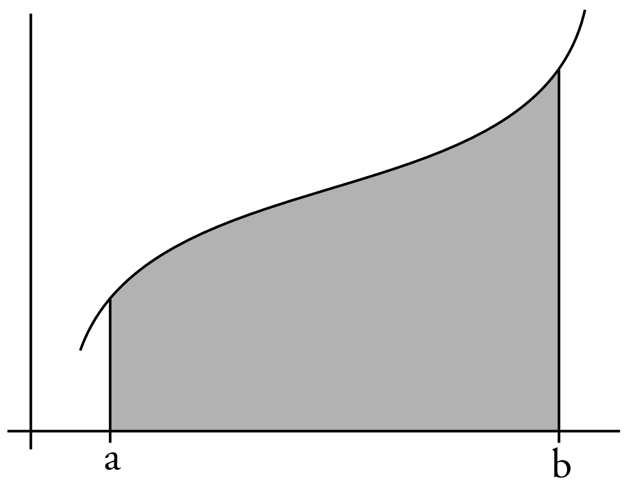

Figure 1: Area under a curve

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.

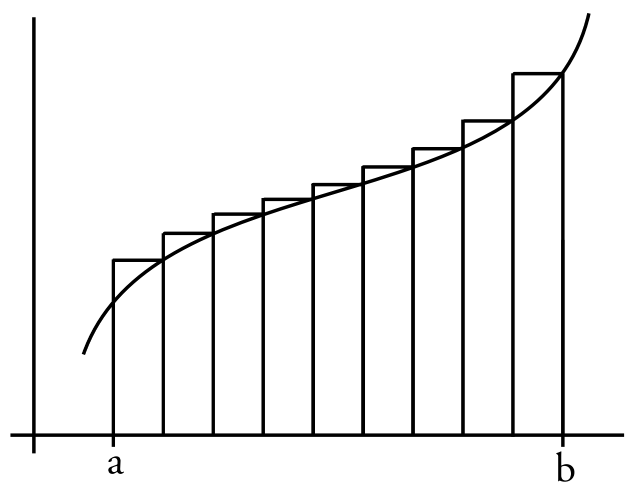

Figure 2: Area under a curve divided into rectangles

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.

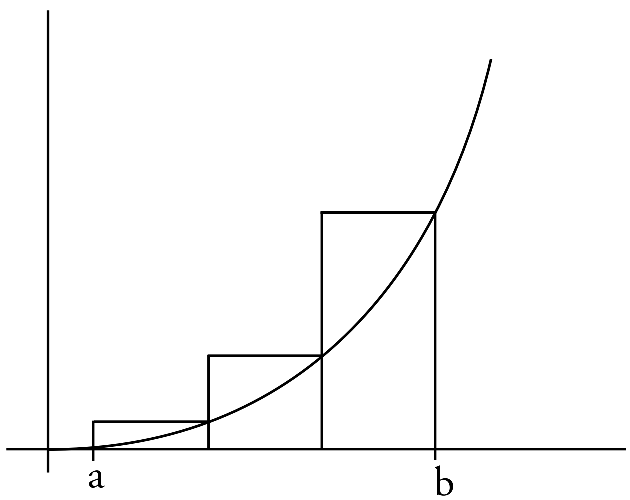

Figure 1: $\int_a^b f(x)dx$ approximated by three rectangles

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

$$\frac{1}{3}n^3 < 1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2.$$

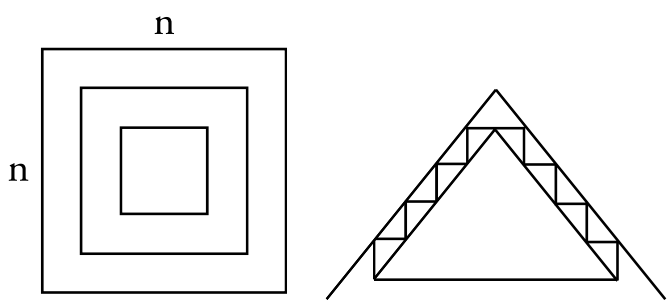

Figure 2: Top and side views of a stair-step pyramid.

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 $\displaystyle\sum_{i=1}^{n} a_i$ 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 $\int_0^b f(x)\,dx$.

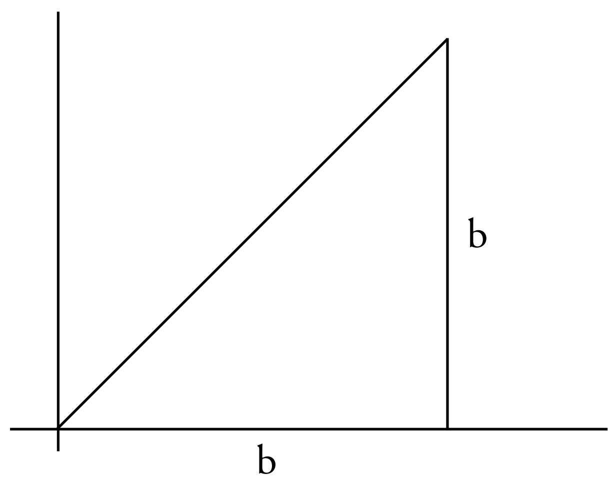

Figure 1: The area under the curve is $\int_0^b xdx$

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.

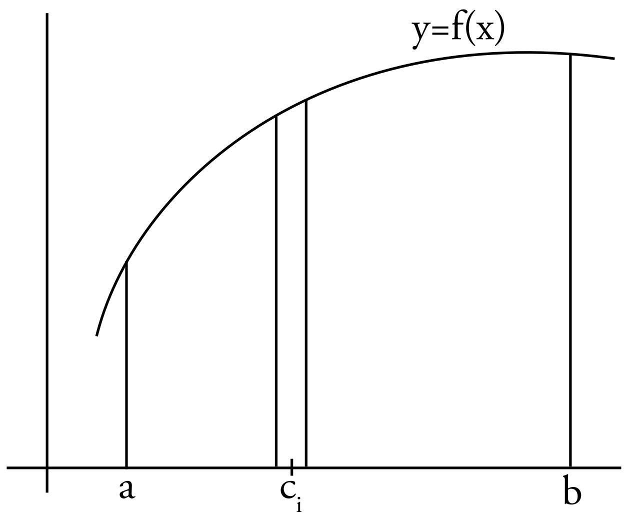

Figure 1: Area under a curve

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 $45^{\mathrm{t}h}$ 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:

$$F(b) - F(a) = \left. F(x)\right|_a^b = \left. F(x)\right|_{x=a}^{x=b}.$$

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:

$$\int_a^b f(x)dx = \left. F(x)\right|_a^b.$$

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 $\sin(x)$.

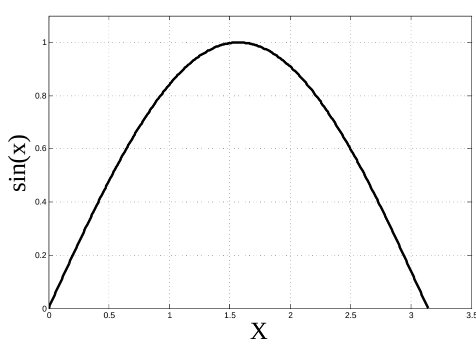

Figure 1: $\sin(x)$ for $0 < x < \pi$

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.

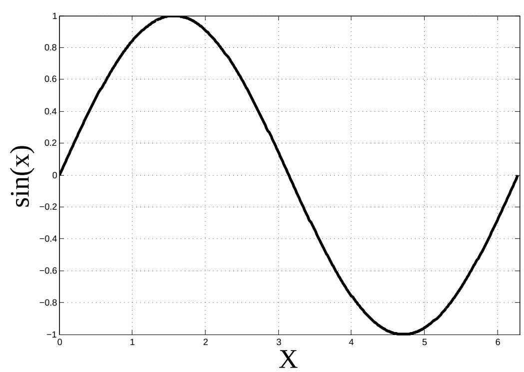

Figure 1: $\sin(x)$ for $0 < x < 2\pi$

$$\int_0^{2\pi} \sin x\, dx = (-\cos x)\Big|_0^{2\pi} = -\cos(2\pi) - (-\cos(0)) = -1 - (-1) = 0$$

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:

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx$$

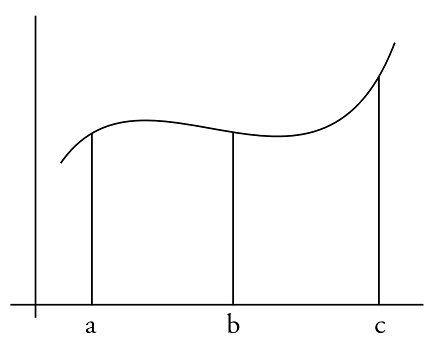

Figure 1: Combining two areas under a curve

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 $f(x) \leq g(x)$ and $a < b$, then $\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$.

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 $e^x \geq 1$ for $x \geq 0$; this is our starting place. We integrate this expression, then follow our noses to get the result we're expecting:

$$\begin{aligned} e^x &\geq 1 \quad (x \geq 0) \\ \int_0^b e^x\,dx &\geq \int_0^b 1\,dx \quad (b \geq 0) \\ e^x\big|_0^b &\geq b \quad \text{(area of rectangle with base } b \text{ and height 1.)} \\ e^b - 1 &\geq b \\ e^b &\geq 1 + b \quad (b \geq 0) \end{aligned}$$

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 $e^b \geq 1 + b$.

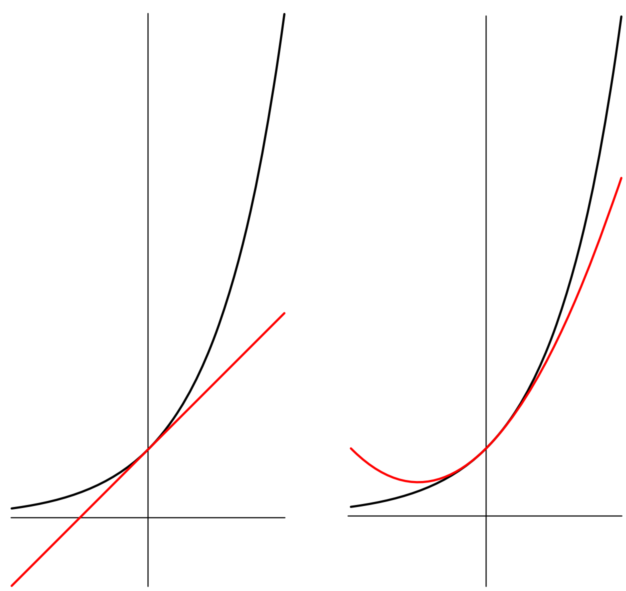

Figure 1: The graphs of e^x (black) compared to $1 + x$ and $1 + x + \frac{x^2}{2}$ (red).

Now we repeat the process starting from the conclusion:

$$\begin{aligned} e^x &\geq 1 + x \quad (x \geq 0) \\ \int_o^b e^x\,dx &\geq \int_0^b (1 + x)\,dx \quad (b \geq 0) \end{aligned}$$

$$
\begin{aligned}
e^b - 1 &\geq \left(x + \frac{x^2}{2} \right) \Bigg|_0^b \\
e^b - 1 &\geq b + \frac{b^2}{2} \\
e^b &\geq 1 + b + \frac{b^2}{2} \quad (b \geq 0)
\end{aligned}
$$

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?

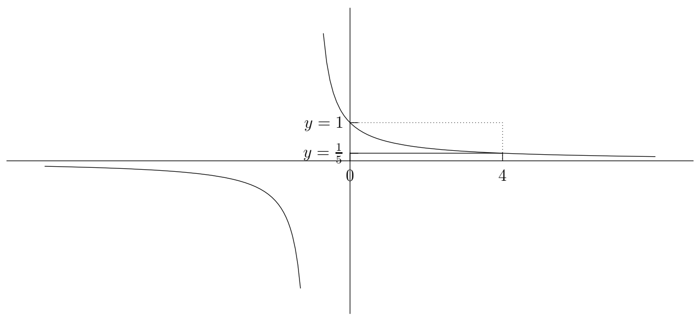

Figure 1: Graph of $F'(x) = \frac{1}{1+x}$.

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.

$$\begin{aligned} F(4) - F(0) &= F'(c)(4-0) \quad \text{(Use the MVT on } F(4)) \\ &= \frac{1}{1+c} \cdot 4 \end{aligned}$$

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.

$$\text{lower Riemann sum} < \int_0^4 \frac{dx}{1+x} < \text{upper Riemann sum}$$