Math 224: Integral Calculus of One Variable Functions
Northwestern University, Lecture Notes

Written by Santiago Cañez

These are notes which provide a basic summary of each lecture for Math 224, "Integral Calculus of One Variable Functions", taught by the author at Northwestern University. The book used as a reference is the 2nd edition of *Essential Calculus: Early Transcendentals* by Stewart. Watch out for typos! Comments and suggestions are welcome.

Contents

Lecture 1: Areas and Distances

Overview. In this course we will continue the study of single-variable calculus, focusing on the notions of *integrals* and *series*. There are two overarching themes behind both of these topics: first the power of using approximations to solve concrete problems, and second the idea of making sense of what it means to add together infinitely many quantities. Indeed, integrals are defined via an approximation procedure but in the end should be thought of as a type of "infinite summation", while series are defined directly as a type of infinite summation but in the end can be used to give good approximations to complicated functions. It's not expected that you know what any of this means at this point, but is an idea we'll highlight again and again.

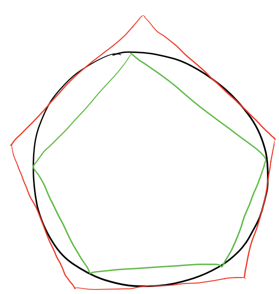
Motivation. In the most basic formulation, integrals are used to compute areas. Of course, certain areas don't require anything fancy to compute—say the area of a triangle or rectangle—but other areas require more care.

For instance, you have likely seen that the area of a unit circle is π, which is a number whose decimal expression looks like

$$\pi = 3.1415926536\ldots.$$

Of course, nowadays we can just have our calculator or compute tell us that this is what π is, but how on earth could we figure this out without having such tools at our disposal? In particular, π was known to the Ancient Greeks—how did they come to know that such a number existed? The key point is that we can actually *approximate* the area of a unit circle using areas of simpler-to-understand figures, and doing so is precisely what lead the Ancient Greeks to discover π.

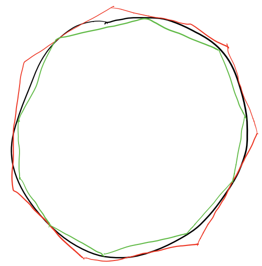
To see how this works, imagine a unit circle surrounded by, say, a pentagon and also enclosing a pentagon:



The areas of these pentagons are possible, although somewhat tedious, to work out by hand, say by dividing the pentagon into five triangles. Given the way in which these shapes are arranged we can see that

$$\text{area of small pentagon} \leq \text{ area of circle} \leq \text{ area of large pentagon}.$$

But now, there is no reason to stop with pentagons! We can do the same thing using figures with more sides:

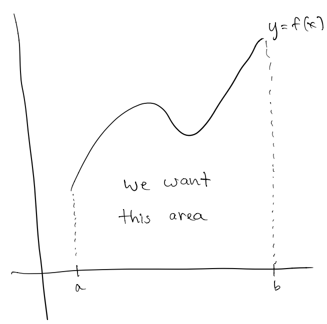


The idea is that as the number of sides increases, the closer the figures come to approximating the circle, and at each step of the way we have:

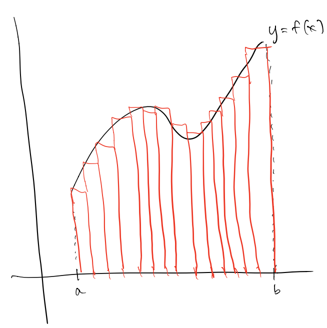
$$\text{area of small } n\text{-sided figure} \leq \text{ area of circle} \leq \text{ area of large } n\text{-sided figure.}$$

As n gets larger and larger, the term on the left gets closer and closer to the actual area of the circle, as does the term on the right. In fact, after taking a *limit* as n goes to infinity, the leftmost term approaches the *same* value as does the rightmost term, and this value is precisely π. Since the area of the circle is always sandwiched between these values, the area of the circle must be this common value π. This limiting process where we approximate the area of the circle using multisided figures is indeed how the Ancient Greeks discovered π, and gives a glimpse of how we can use approximations to derive concrete information.

Areas via rectangles. The type of area we will be interested in for now is the area of a region lying under the graph of a function:

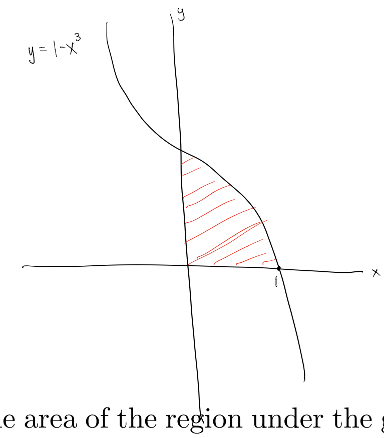


We will define integrals in more detail later on, but just to introduce the terminology now, this area is essentially what the so-called *integral* of f from a to b will give us. The key idea is that we can determine this area through some limiting process where we *approximate* it using *rectangles*:



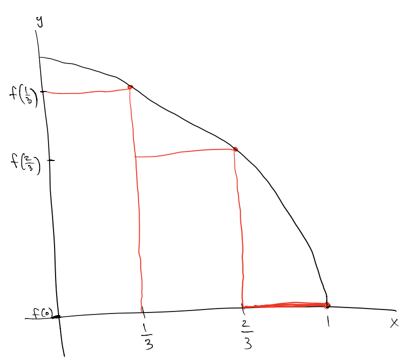
We will talk about this "limiting process" next time, and so will focus today on the approximation aspect, which is easiest to do by working through an example.

Example. Consider the function $f(x) = 1 - x^3$, whose graph looks like:



We want to eventually compute the area of the region under the graph of this function and between $x = 0$ and $x = 1$, which is the portion shaded in red above. Let us see how we can approximate this area using rectangles, which are simple geometric figures whose areas are possible to write down explicitly.

Take the entire interval between $x = 0$ and $x = 1$ on the x-axis, and divide it into 3 *subintervals* of equal length, so in this case the intervals from $x = 0$ to $x = 1/3$, $x = 1/3$ to $x = 2/3$, and $x = 2/3$ to $x = 1$. We use each of these subintervals as the *base* of a rectangle whose height is given by the value of the function f at some point within that subinterval. For instance, we first use the values at the *right endpoint* of each subinterval as the height:



so the first rectangle has height given by the value of $f(1/3)$, the second height given by $f(2/3)$, and the third height given by $f(1)$. Note that in this case it just so happens that $f(1) = 0$, so the final rectangle has no height and is thus represented by a "flat" line between $x = 2/3$ and $x = 1$. Nonetheless, we'll still include this (not really a) rectangle in our equations to illustrate the general process.

For each rectangle, its area is given by

$$f(\text{right endpoint of subinterval}) \cdot (\text{length of subinterval}).$$

In this case we get

$$f\left(\frac{1}{3}\right)\frac{1}{3} = \left(1 - \frac{1}{3^3}\right)\frac{1}{3} \qquad f\left(\frac{2}{3}\right)\frac{1}{3} = \left(1 - \frac{2^3}{3^3}\right)\frac{1}{3} \qquad f(1)\frac{1}{3} = (1-1)\frac{1}{3}$$

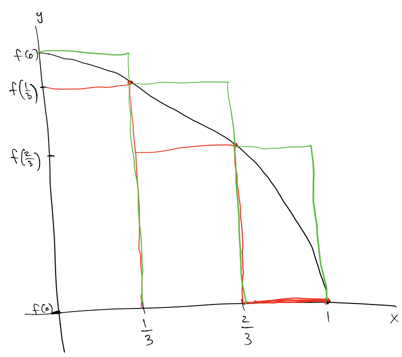
as the areas of the three rectangles. The sum of these areas is meant to approximate the area of the region we want, and this sum is:

$$\left(1-\frac{1}{3^3}\right)\frac{1}{3}+\left(1-\frac{2^3}{3^3}\right)\frac{1}{3}+(1-1)\frac{1}{3}, \text{ which is approximately } 0.556.$$

Note that in this case, this value *underestimates* the area we want since the rectangles in question are fully enclosed without our region but don't cover all of it; this means that

$$0.556 \leq \text{area we want}.$$

Now, there was nothing special about using right endpoints to give the heights of our rectangles, and we could just as well have used left endpoints instead. In this case we get rectangles which look like (in green):



with the previous rectangles still in red. In this case, the sum of the areas of the three green rectangles is:

$$f(0)\frac{1}{3}+f\left(\frac{1}{3}\right)\frac{1}{3}+f\left(\frac{2}{3}\right)\frac{1}{3}=(1-0)\frac{1}{3}+\left(1-\frac{1}{3^3}\right)\frac{1}{3}+\left(1-\frac{2^3}{3^3}\right)\frac{1}{3},$$

which is approximately 0.889. In this case, this value *overestimates* the area in question since the green rectangles always go above the region in question; this gives

$$\text{area we want} \leq 0.889,$$

which together with the previous inequality gives

$$0.556 \leq \text{area we want} \leq 0.889.$$

Warning. Because of the decreasing nature of the function $f(x)=1-x^3$, it turns out that in this case using left endpoints to give heights of rectangles will always lead to an overestimation of the area and using right endpoints will always lead to an underestimation. However, this does not happens for all functions; for instance, with an increasing function the opposite would be true: left endpoints would underestimate and right endpoints will overestimate. For functions which are neither increasing nor decreasing, other things can happen.

Taking a limit. Going back to our example, the idea is that as we carry out the same procedure only using more and more rectangles, obtained by splitting our original interval from $x=0$ to

$x = 1$ into more and more pieces, the values we get will provide better and better approximations to the area we want. For instance, if we had used $n = 4$ subintervals:

$$x = 0 \text{ to } x = \frac{1}{4}, \quad x = \frac{1}{4} \text{ to } x = \frac{2}{4}, \quad x = \frac{2}{4} \text{ to } x = \frac{3}{4}, \quad x = \frac{3}{4} \text{ to } x = 1,$$

using left endpoints to give heights leads to rectangles whose areas add up to:

$$f(0)\frac{1}{4} + f\left(\frac{1}{4}\right)\frac{1}{4} + f\left(\frac{2}{4}\right)\frac{1}{4} + f\left(\frac{3}{4}\right)\frac{1}{4} = (1-0)\frac{1}{4} + \left(1 - \frac{1}{4^3}\right)\frac{1}{4} + \left(1 - \frac{2^3}{4^3}\right)\frac{1}{4} + \left(1 - \frac{3^4}{4^4}\right)\frac{1}{4},$$

which is approximately 0.859. The actual area of the region in question is 0.75 (we'll see how to definitely compute this later), so the 0.859 value we got using left endpoints and 4 subintervals is a better approximation than the 0.889 value we got using only 3 subintervals. Using right endpoints and 4 subintervals gives an approximate value of 0.609, which is closer to the actual 0.75 than the 0.556 from before. To get the actual area we would continue this process, taking n subintervals with n getting larger and larger, eventually taking a limit as n goes to infinity.

This process works (!), but the problem is that limit expression obtained in this final step is very difficult to work with in practice, and it is only in certain circumstances that it becomes directly computable. Nonetheless, this limiting process is what we'll use to define integrals soon enough, but the upshot is that this is NOT the method we'll use to actually compute integrals. Magically, it turns out that there is an alternate much more direct way of computing values of integrals using *antiderivatives*, and it is this discovery which made Isaac Newton famous in the 17th century. This is the content of the *Fundamental Theorem of Calculus*, which we'll talk about next week. For now, we are simply highlighting how we can use rectangles to approximate areas, foreshadowing other approximation techniques we'll see later on.

Distance. We finish with what seems to be totally unrelated question to the one we were considering above, but turns out to be a reflection of the same idea. Suppose Usain Bolt (I've still got the Olympics on my mind!) is running and someone measures his speed (say in meters per second) every 5 seconds for 20 seconds:

time	5	10	15	20
speed	10	15	13	8

The problem is to estimate how much distance Usain traveled over these 20 seconds. Knowing that

$$\text{distance} \ = \ \text{speed times time},$$

we can estimate that over the first 5 seconds he traveled

$$10 \cdot 5 = 50 \text{ meters},$$

after the next 5 seconds he traveled

$$15 \cdot 5 = 75 \text{ meters},$$

and so on. Overall he covered a distance of approximately

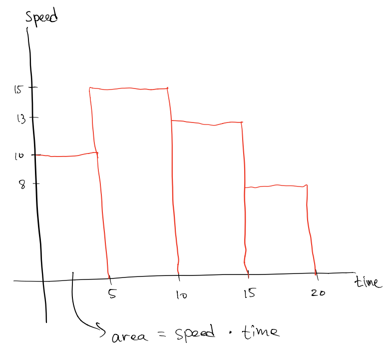
$$10 \cdot 5 + 15 \cdot 5 + 13 \cdot 5 + 8 \cdot 5 = 230 \text{ meters}.$$

(So, this seems unrealstic: Usain Bolt's world record time for 200 meters is just under 20 seconds, and it seems unlikely he could travel the remaining 30 meters in under 1 second as our numbers would suggest, but oh well.)

The point to recognize is that the type of sum we ended up with above:

$$10 \cdot 5 + 15 \cdot 5 + 13 \cdot 5 + 8 \cdot 5$$

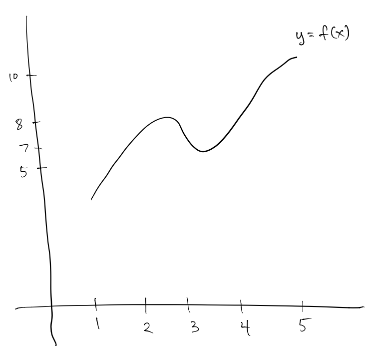
is similar to the type of sum we saw previously when approximating areas using rectangles. Indeed, if we set up a pair of axes where the horizontal axis is time and the vertical axis speed:



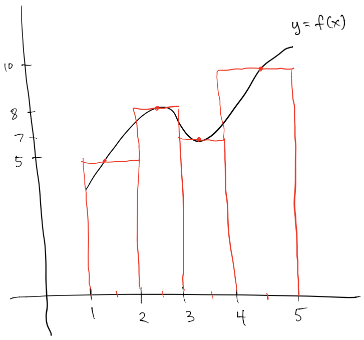
the sum above can be interpreted as the sum of the areas of the rectangles shown here. The upshot is that this problem, although not directly phrased in terms of areas, can in fact be interpreted in such a way, and suggests that *integration* in general has many applications which go way beyond area problems. Indeed, we'll see that distance can indeed be obtained by integrating speed, and the sum we used above to approximate such a distance is at the core of the reason why.

Lecture 2: The Definite Integral

Warm-Up. We estimate the area under the graph of the function drawn below between $x = 1$ and $x = 5$:



We do so using rectangles obtained by dividing the interval from $x = 1$ to $x = 5$ into four pieces of equal length, and using *midpoints* to determine the heights. Thus we use the following rectangles:



The sum of the areas of these rectangles is

$$f(1.5)(1) + f(2.5)(1) + f(3.5)(1) + f(4.5)(1),$$

where as stated we use the value of f at the midpoint of each subinterval as the height. Based on the given graph, these values are:

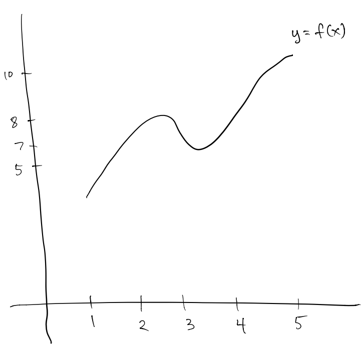
$$f(1.5) = 5 \qquad f(2.5) = 8 \qquad f(3.5) = 7 \qquad f(4.5) = 10,$$

so the sum of the areas of the rectangles is:

$$5(1) + 8(1) + 7(1) + 10(1) = 30.$$

Thus we approximate the area under the graph of the given function between $x = 1$ and $x = 5$ as being close to 30. Using more rectangles should give better estimates.

A modification. Here's a seemingly different question. Suppose water is being added to a pool, and that the rate at which the water flows in at a given time is given by the function:



Now we interpret the x axis as time (say in seconds) and the vertical axis as a rate. So, the value $f(x)$ gives the rate at which water flows in at time x. We want to estimate the total amount of water which went into the pool between $x = 1$ and $x = 5$ seconds.

The point is that if water flows in at a constant rate over a period of some number of seconds, the amount which flowed in over that time period is

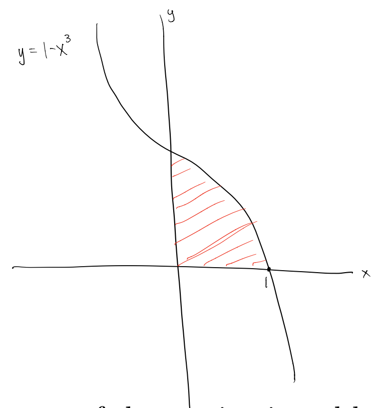
$$\text{amount} = \text{rate} \cdot \text{time}.$$

If we use the time intervals $x = 1$ second to $x = 2$ seconds, $x = 2$ seconds to $x = 3$ seconds, and so on, and we use the half-way points in each such subinterval as the estimate for the rate over that corresponding subinterval, then we estimate the total amount of water which flows in between 1 and 5 seconds to be:

$$f(1.5)(1) + f(2.5)(1) + f(3.5)(1) + f(4.5)(1) = 5(1) + 8(1) + 7(1) + 10(1) = 30.$$

The observation of course is that the sum on the left is the *same* sum we used in the Warm-Up area problem, the connection being that in both cases we are adding up various products. As mentioned last time, this suggests that integrals have wide applications going way beyond computing areas alone.

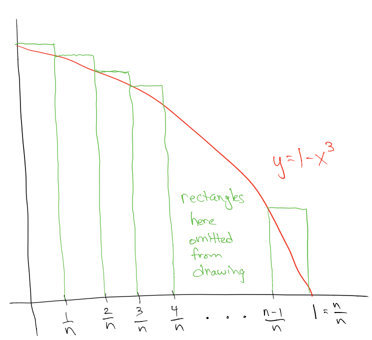
An explicit area. Let us now try to compute a precise area, no longer approximate areas. Indeed, let's return to the function $f(x) = 1 - x^3$ we used last time:



and actually work out the concrete area of the portion in red between $x = 0$ and $x = 1$. The idea is one we've mentioned: we approximate this using rectangles, but then take a limit as the number of rectangles grows larger and larger. For some number n, we divide the interval from $x = 0$ to $x = 1$ into n subintervals of equal length. Since our entire interval has length 1, this means that each subinterval will have length $\frac{1}{n}$, and so the subintervals we use are:

$$x = 0 \text{ to } x = \frac{1}{n}, \quad x = \frac{1}{n} \text{ to } x = \frac{2}{n}, \quad x = \frac{2}{n} \text{ to } x = \frac{3}{n}, \quad \text{and so on,}$$

with the final subinterval being the one from $x = \frac{n-1}{n}$ to $x = \frac{n}{n} = 1$. We will use left endpoints when determining the heights of rectangles, so the rectangles we use look like:



The left endpoints of these given subintervals are $0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}$, so the sum of the areas of the rectangles is

$$f(0)\frac{1}{n} + f\left(\frac{1}{n}\right)\frac{1}{n} + f\left(\frac{2}{n}\right)\frac{1}{n} + \cdots + f\left(\frac{n-1}{n}\right)\frac{1}{n},$$

where the $\cdots$ indicate a bunch of intermediate terms.

Riemann sums. The type of expression we ended up with above:

$$f(0)\frac{1}{n} + f\left(\frac{1}{n}\right)\frac{1}{n} + f\left(\frac{2}{n}\right)\frac{1}{n} + \cdots + f\left(\frac{n-1}{n}\right)\frac{1}{n}$$

is known as a *Riemann sum* for the function f. Geometrically, a Riemann sum is nothing but the sum of areas of a bunch of triangles. Now, Riemann sums come in all shapes and flavors, depending on which specific subintervals are used and which "sample points" we use to give the heights of the rectangles in question. For instance, in general there is no reason why we couldn't use subintervals of unequal length when estimating areas, and there is no reason why we couldn't use a left endpoint for one subinterval, a midpoint for another, and some other random point from another. Any such sum obtained by such choices is still a Riemann sums, it's just that certain Riemann sums lead to nicer expression.

You can check the book for more formal notation and terminology (such as the notion of a *partition*) associated with Riemann sums, but we won't dwell too much on such things. After all, we'll soon get to a point where we can compute these areas without considering Riemann sums at all, so for now we're just getting a glimpse as to how Riemann sums work.

Definite integrals. The idea, then, is to consider Riemann sums as the width of the rectangles used becomes smaller and smaller, which in turn will produce more and more rectangles. As the widths approach zero (or correspondingly the number of rectangles goes off to infinity), the values obtained give better and better approximations to the area under the graph of the function, so that this area is obtained as a *limit* of such Riemann sums.

In general, given a function f and an interval from $x = a$ to $x = b$, we define the *definite integral* of f from $x = a$ to $x = b$ to be the limit of Riemann sums for f over the interval $x = a$ to $x = b$ as the widths of the subintervals used goes to zero:

$$\int_a^b f(x)\,dx = \text{ limit of Riemann sums as widths go to zero.}$$

The notation on the left is our notation for such a definite integral, where the lower and upper "bounds" on the integral symbol (the elongated S symbol at the beginning) indicate the interval $x = a$ to $x = b$ we're looking at. For now the "dx" should be viewed as simply part of the notation, but we'll see what it is meant to stand for soon. The book has this definition written out in more formal notation, but the core idea is what I wrote above. Geometrically, such integrals precisely compute the areas we've been considering:

$$\int_a^b f(x)\,dx = \text{"area" between the graph of } f \text{ and the } x\text{-axis between } x = a \text{ and } x = b.$$

We'll see later what I mean by "area". (The point is that the value can actually be negative.)

Back to explicit area. Returning to our example, we're looking at the Riemann sum

$$f(0)\frac{1}{n} + f\left(\frac{1}{n}\right)\frac{1}{n} + f\left(\frac{2}{n}\right)\frac{1}{n} + \cdots + f\left(\frac{n-1}{n}\right)\frac{1}{n}.$$

We have a nice way of expressing such sums using more compact notation, so called "sigma" notation:

$$\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right)\frac{1}{n}.$$

The $\sum$ symbol is the Greek letter "sigma", and should be thought of as an "S" for "Sum". This notation indicates the sum of the terms

$$f\left(\frac{k}{n}\right)\frac{1}{n}$$

as k varies from $k = 0$ to $k = n - 1$, increasing the value of k by 1 at each step. So, the first term when $k = 0$ would be

$$f(0)\frac{1}{n},$$

the second term when $k = 1$ is

$$f\left(\frac{1}{n}\right)\frac{1}{n},$$

the term when $k = 2$ is

$$f\left(\frac{2}{n}\right)\frac{1}{n},$$

and so on, with the final term when $k = n - 1$ being

$$f\left(\frac{n-1}{n}\right)\frac{1}{n}.$$

The sum of all these is indeed the Riemann sum we were considering before:

$$\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right)\frac{1}{n} = f(0)\frac{1}{n} + f\left(\frac{1}{n}\right)\frac{1}{n} + f\left(\frac{2}{n}\right)\frac{1}{n} + \cdots + f\left(\frac{n-1}{n}\right)\frac{1}{n}.$$

Sigma notation is often simpler to work with than the expression on the right with the ambiguous-looking $\cdots$ occurring in the middle.

Now, the actual function we're looking at is $f(x) = 1 - x^3$, so

$$f\left(\frac{k}{n}\right) = 1 - \frac{k^3}{n^3}.$$

Thus our Riemann sum is actually:

$$\sum_{k=0}^{n-1}\left(1 - \frac{k^3}{n^3}\right)\frac{1}{n}.$$

As n goes off to infinity, the widths of our subintervals goes to zero, and so the definition of the definite integral we gave above would say that:

$$\int_0^1 (1 - x^3)\,dx = \lim_{n\to\infty}\sum_{k=0}^{n-1}\left(1 - \frac{k^3}{n^3}\right)\frac{1}{n}.$$

Our goal is now to figure what this sum actually is, so that we can then actually determine this limit, and hence the area we want, precisely.

We can write our Riemann sum as

$$\sum_{k=0}^{n-1}\left(1 - \frac{k^3}{n^3}\right)\frac{1}{n} = \sum_{k=0}^{n-1}\left(\frac{1}{n} - \frac{k^3}{n^4}\right)$$

after we distribute the $\frac{1}{n}$ term to both terms in the parenthesis. Recall that this sigma notation denotes the sum:

$$\left(\frac{1}{n}-\frac{0^3}{n^4}\right)+\left(\frac{1}{n}-\frac{1^3}{n^4}\right)+\left(\frac{1}{n}-\frac{2^3}{n^4}\right)+\cdots+\left(\frac{1}{n}-\frac{(n-1)^3}{n^4}\right).$$

Now, all together in this sum we have n copies of $\frac{1}{n}$, one for each of the subintervals we had. Grouping these terms together gives $n\frac{1}{n}=1$, and grouping the remaining terms together gives

$$-\frac{0^3}{n^4}-\frac{1^3}{n^4}-\frac{2^3}{n^4}-\cdots-\frac{(n-1)^3}{n^4}=-\frac{1}{n^4}(0^3+1^3+2^3+\cdots+(n-1)^3).$$

Thus our Riemann sum simplifies to:

$$\sum_{k=0}^{n-1}\left(1-\frac{k^3}{n^3}\right)\frac{1}{n}=\sum_{k=0}^{n-1}\left(\frac{1}{n}-\frac{k^3}{n^4}\right)=1-\frac{1}{n^4}(0^3+1^3+2^3+\cdots+(n-1)^3)).$$

Now, it is a fact that

$$0^3+1^3+2^3+\cdots+(n-1)^3=\sum_{k=0}^{n-1}k^3=\frac{(n-1)^2n^2}{4}.$$

This is *not* at all obvious and requires a good amount of work to justify; this is the type of thing which you would have to be given as a fact if you were to be required to use it since no one in their right mind (except for mathematicians) would know this off the top of their head. Ask in office hours if you are *really* interested in seeing where this comes from. The upshot is that we can finally write our Riemann sum as:

$$1-\frac{1}{n^4}(0^3+1^3+2^3+\cdots+(n-1)^3))=1-\frac{1}{n^4}\frac{(n-1)^2n^2}{4}.$$

After rewriting this expression so that everything uses the same denominator and simplifying, we get:

$$1-\frac{1}{n^4}(0^3+1^3+2^3+\cdots+(n-1)^3))=\frac{4n^4-(n-1)^2n^2}{4n^4}=\frac{3n^4+2n^3-n^2}{4n^4},$$

so that finally after all this work we get that:

$$\int_0^1(1-x^3)\,dx=\lim_{n\to\infty}\frac{3n^4+2n^3-n^2}{4n^4}.$$

WHEW!!! That was a lot of work! But fear not, the point is that this *should* seem tedious and *should* make you hope that there is a simpler way of computing integrals, which there is (!), as we'll talk about next week. But, going through this type of thing at least one is good for getting a sense of where things actually come from, and for making you appreciate just how important the *Fundamental Theorem of Calculus* we'll look at next week actually is.

The final limit obtained is now one we can compute using ideas from Math 220. For instance, by dividing the numerator and denominator both by n^4 we can rewrite this limit as

$$\lim_{n\to\infty}\frac{3n^4+2n^3-n^2}{4n^4}=\lim_{n\to\infty}\frac{3+\frac{2}{n}-\frac{1}{n^2}}{4},$$

which is thus equal to $\frac{3}{4}$ since the $\frac{2}{n}$ and $\frac{1}{n^2}$ terms both go to zero as n goes to infinity. Thus, we finally get our desired value, the area of the region we've been considering is

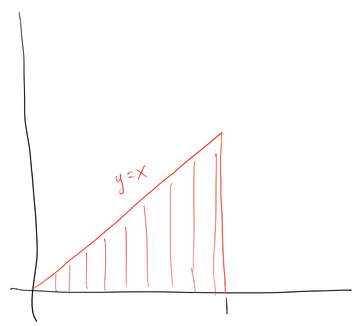
$$\int_0^1 (1 - x^3)\, dx = \frac{3}{4}.$$

We mentioned this value (or rather, 0.75) in passing last time, when assessing how good the estimates we found for this area actually were. Now we know where it came from.

Integrals via geometry. The fact that integrals can be interpreted as areas makes certain integrals easily computable now, using what we already know about geometry. For instance, the integral

$$\int_0^1 x\, dx$$

gives the area of the following triangular region, below the line $y = x$ between $x = 0$ and $x = 1$:



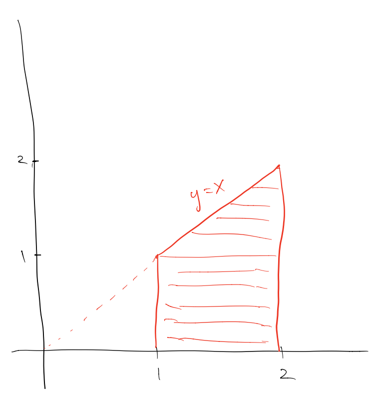
This triangle has area $\frac{1}{2}$, so

$$\int_0^1 x\, dx = \frac{1}{2}.$$

The integral

$$\int_1^2 x\, dx$$

gives the area of the region:



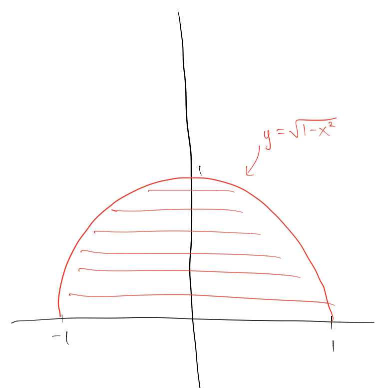
By thinking of this as a square of height 1 with a triangle on top we can see that this region has area $\frac{3}{2}$, so

$$\int_1^2 x\,dx = \frac{3}{2}.$$

Finally, consider the integral

$$\int_{-1}^1 \sqrt{1-x^2}\,dx.$$

The curve $y = \sqrt{1-x^2}$ describes the top half of the unit circle, which we can see by squaring both sides to get $y^2 = 1 - x^2$ and then rearranging to get $x^2 + y^2 = 1$. We get the to half since we are taking the *positive* square root of $1 - x^2$; the bottom half would have equation $y = -\sqrt{1-x^2}$. Thus the integral above gives the area of the region:



This is half the region enclosed by the entire unit circle, so since the unit circle has area π, this particular region has area $\frac{\pi}{2}$, so

$$\int_{-1}^1 \sqrt{1-x^2}\,dx = \frac{\pi}{2}.$$

Of course, not all integrals are so easily computable, since the areas in question are not always ones of well-known geometric figures. We'll see next week the "real" way to compute such integrals.

Lecture 3: Evaluating Definite Integrals

Warm-Up. We determine the value of the limit

$$\lim_{n\to\infty} \sum_{k=1}^n \frac{1}{n}\sqrt{1-\frac{k^2}{n^2}}$$

by interpreting it as an area via an integral. (Note that the original version I gave in class used $\frac{2}{n}$ instead of $\frac{1}{n}$ and $\frac{2k}{n}$ instead of $\frac{k}{n}$. I'll remind you below why I changed this to the version given here.) The point is that this limit is pretty much impossible to compute directly, but by recognizing what it is actually mean to represent, it becomes much more manageable. The types of limits of Riemann sums which define integrals look like

$$\lim_{n\to\infty} \sum_{k=1}^n f(\text{some value})(\text{width}).$$

So, in order to interpret the given limit as such an expression, we must determine what is the "width" term, what is the function being considered, and what is the interval being integrated over.

The $\frac{1}{n}$ term in our case plays the role of "width". The $\frac{k}{n}$ terms represent sample points from the subintervals being considered. In this case, as k increases from $k=1$ to $k=n$, these $\frac{k}{n}$ same points run through the values

$$\frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \ldots, \frac{n}{n} = 1.$$

Thus we are looking at the interval $[0,1]$ being divided into n pieces of equal length, using *right endpoints* as sample points. (Using left endpoints would have corresponded to $0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}$, which would have arisen if the sum in question ran from $k=0$ to $k=n-1$.)

Since the $\frac{1}{n}$ in the original summation thus corresponds to the lengths of the subintervals obtained by breaking $[0,1]$ into n pieces of equal length, the remaining $\sqrt{1-\frac{k^2}{n^2}}$ term must correspond to the

$$f(\text{some value})$$

piece which gives the height of a rectangle making up our Riemann sum. Thus, the function being considered is

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

and note that evaluating this at the given right endpoints $\frac{k}{n}$ indeed gives

$$f\left(\frac{k}{n}\right) = \sqrt{1-\frac{k^2}{n^2}}$$

as we need. (The reason why using $\frac{2}{n}$ and $\frac{2k}{n}$ as I originally had is bad is that this would correspond to the interval $[0,2]$, since when $k=n$ the final right endpoints $\frac{2k}{n}$ would be 2. However, the function $f(x) = \sqrt{1-x^2}$ is not defined for $x > 1$ since these values give a negative term under the square root. This is why I modified the original sum so that only the interval $[0,1]$ would be considered.) The overall point is that, for the function $f(x) = \sqrt{1-x^2}$, the expression

$$\frac{1}{n}\sqrt{1-\frac{k^2}{n^2}} \text{ corresponds to (width)(}f\text{ evaluated at right endpoints),}$$

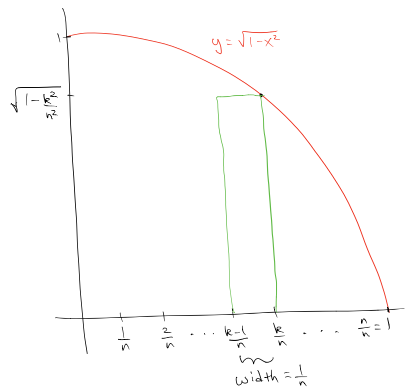
so that the sum

$$\sum_{k=1}^{n} \frac{1}{n}\sqrt{1-\frac{k^2}{n^2}}$$

is a Riemann sum for $f(x) = \sqrt{1-x^2}$ over the interval $[0,1]$. Thus, the limit in question is precisely the integral of this function over the interval $[0,1]$:

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{n}\sqrt{1-\frac{k^2}{n^2}} = \int_0^1 \sqrt{1-x^2}\, dx.$$

The upshot is that this is an integral we can compute geometrically, based on its interpretation as an area. Indeed, we saw last time that the graph of $f(x) = \sqrt{1-x^2}$ is the upper-half of the unit circle, so here we are considering the area under this upper-circle from $x=0$ to $x=1$, which is thus a quarter of a circle overall:



(The area of the rectangle in green is given by one of the terms $\frac{1}{n}\sqrt{1-\frac{k^2}{n^2}}$ in our Riemann sum.) The full unit circle has area π, so our region has area $\frac{\pi}{4}$. Hence

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{n}\sqrt{1-\frac{k^2}{n^2}} = \int_0^1 \sqrt{1-x^2}\,dx = \frac{\pi}{4}$$

is the desired value of the limit on the left.

Net area. The value of the integral

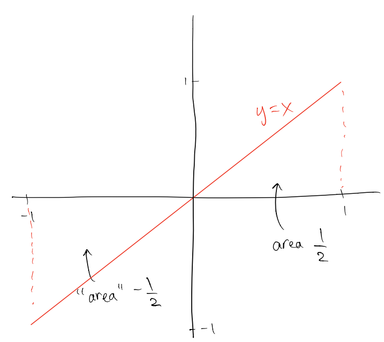
$$\int_{-1}^{1} x\,dx$$

is zero, which seems to go against the idea that integrals compute areas. The point of clarification we now make is that it is *not* literally true that integrals compute areas, but rather what they actually compute is "net area". The key observation is that "areas" under the x-axis are actually counted as being negative, so an integral can in general have positive and negative area contributions, and what we get is a sum of these contributions. Indeed, the fact that "areas" below the x-axis are counted as being negative comes directly from the Riemann sum approach: in an "area" term given by

$$f(\text{sample point})(\text{width of subinterval}),$$

the "height" f(sample point) can be negative depending on the function f, so that the product above is not the actual area of some rectangle but rather the negative of this area.

In the example above, the region between the graph of $f(x) = x$ and the x-axis between $x = -1$ and $x = 1$ looks like:



This region consists of two triangles of equal area, only that one is counted as positive and one as negative, so that adding them together gives zero:

$$\int_{-1}^{1} x\,dx = 0.$$

Another way to phrase this is as follows. Splitting the interval entire interval $[-1,1]$ into two pieces $[-1,0]$ and $[0,1]$ gives a corresponding splitting of intervals:

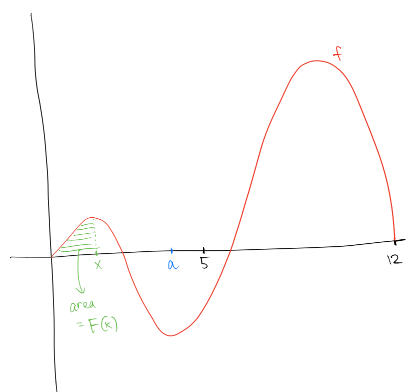
$$\int_{-1}^{1} x\,dx = \int_{-1}^{0} x\,dx + \int_{0}^{1} x\,dx.$$

Each integral on the right measures the net area of one of the triangles in the picture we're looking at, only that the first integral on the right gives negative the area (which is $-\frac{1}{2}$) and the second gives positive the area (which is $\frac{1}{2}$). Thus, the first integral is precisely the negative of the second, so their sum is zero:

$$\int_{-1}^{1} x\,dx = \int_{-1}^{0} x\,dx + \int_{0}^{1} x\,dx = -\frac{1}{2} + \frac{1}{2} = 0.$$

To summarize, we now have the true geometric meaning behind an integral: $\int_a^b f(x)\,dx$ measures the *net area* of the region between the x-axis and the graph of f between $x = a$ and $x = b$, where areas below the x-axis count as negative and areas above the x-axis count as positive.

Area function. Suppose f is the function whose graph is drawn below in red:



Define a new function F by setting the value of F at x to be

$$F(x) = \int_0^x f(t)\,dt.$$

In other words, based on our geometric interpretation of this integral, $F(x)$ gives the net area between the graph of f and the x-axis from 0 to whatever value x is, where x is allowed to vary. Thus, F measures net area up to some given variable point. For instance, for the given green point x, the area of the region shaded in green is $F(x)$.

A word about the notation. Note that in the definition of F the integral in question uses t as the variable, so that we have $f(t)$ and dt. We should not use x here since we are already using x to denote the variable upper bound on the integral. The point is that x and t denote two separate things: x the point which tells how far to integrate to, and t the variable of integration itself. Note that the fact we use t to denote the variable of integration is irrelevant, and we could just have easily used any other letter or symbol instead. In other words,

$$\int_a^b f(t)\,dt,\ \int_a^b f(u)\,du,\ \text{and}\ \int_a^b f(☺)\,d☺$$

all mean the same thing, as long as the variable used in the function f and the variable used in the "d" part at the end match up.

Now we ask some questions about the function F. For instance, is $F(5) > 0$? Is $F(12) > 0$? Are there any values of x where $F(x) = 0$? First, $F(5)$ gives the net area between the graph of f and the x-axis up to 5, which has a positive contribution coming from the first "hump" and then a negative contribution. The area which gives the negative contribution seems to be larger than the area which gives the positive contribution, so $F(5)$ will be negative. Next, $F(12)$ gives the net area all the way to 12, which has two positive contributions and one negative contribution. Overall, from eyeballing the picture, it seems that the positive contribution areas are larger than the negative area contribution, so $F(12)$ is positive.

Finally, note that of course $F(0) = 0$, since $F(0)$ measures the area from 0 to 0, and there is no such area there. (Or, you can think of this as saying that the area of a single "point" is zero.) However, there is another value a bit before 5 indicated by the blue a at which $F(a)$ will also be zero. Indeed, the net area up to this point has positive and negative contributions, and I've drawn (or at leat tried to) the point a at the point where these two effects would cancel each other out, meaning that area of the first hump above the horizontal axis should be the same as the area of the portion below the horizontal axis from the point where the graph intercepts that axis up to a. There will actually be another point (not drawn) a bit after the graph crosses the horizontal axis and goes positive again at which $F(x)$ will again be zero. We'll come back and see what is special about this type of "area function" next time.

Evaluation Theorem. We now come to what the book refers to as the *Evaluation Theorem*, and which most other sources refer to as the *Fundamental Theorem of Calculus*. (Our book eventually uses "Fundamental Theorem of Calculus, Part II" to refer to the Evaluation Theorem, to contrast it with a very much related fact which also goes by the name "Fundamental Theorem of Calculus". We'll get to this next time.) This is the fact which tells us that integrals can be computed in a much simpler way than having to resort to the limit of Riemann sums definition, which is very difficult to work with in general. The surprising (or "fundamental") fact is that evaluating an integral amounts to finding an *antiderivative*.

The statement is that if F is an antiderivative of f, meaning F is a function whose derivative is f (i.e. $F'(x) = f(x)$ for all x), then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Thus, in order to compute an integral, we first find an antiderivative of the function being integrated, and then we evaluate this antiderivative at our bounds and subtract. We'll see antiderivatives are not always straightforward to find, but for the most part they are simpler to find than the alternate approach of trying to compute some limit of Riemann sums explicitly. The rest of the "integration" half of the course will essentially be devoted to techniques used to find antiderivatives. Notation-wise, the difference $F(b) - F(a)$ showing up above is normally shortened to

$$F(b) - F(a) = F(x)\Big|_a^b \quad \left(\text{ or } F(x)\Big]_a^b \text{ as the book uses}\right),$$

so that the Evaluation Theorem can be written as

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b, \text{ where } F'(x) = f(x).$$

I referred above to the fact that the Evaluation Theorem is in many ways quite surprising, since the definition of an integral via rectangles seems to have nothing to do with derivatives, let alone antiderivatives. The fact that the process of integration is so directly related to the process of differentiation was a *major* discovery in the history of mathematics and is part of what made Isaac Newton famous in the 1600's. People had been working with integral-like objects since the time of the Ancient Greeks, but before Newton (and a mathematician named Leibniz) came along, there only existed a hodge-podge of various techniques for actually evaluating such things, which only worked in certain situations. Newton and Leibniz's realization (and reason why the "Fundamental Theorem of Calculus" is referred to as "fundamental") was that all of these techniques could be encoded more efficiently via anti-differentiation, which was truly a magnificent accomplishment. If nothing else, I hope you come to appreciate how much "simpler" it is to compute integrals in this manner than it is via the Riemann sum definition. ("Simpler" here is used in a relative sense; antiderivatives can still be tricky to compute in general.)

Example 1. We return to some examples we saw previously. Consider

$$\int_0^1 x\,dx.$$

According to the Evaluation Theorem, to compute this integral all we need to do is come up with a function whose derivative is x. Thinking back to what we know about derivatives, we know that in order to get x after differentiation we should start with something like x^2. However, the derivative of x^2 is $2x$, whereas we just want to end up with x. To deal with the extra factor of 2 at the front of $2x$, we should multiply our candidate antiderivative of x^2 by $\frac{1}{2}$ to get

$$\frac{1}{2}x^2.$$

The derivative of this function is indeed just x, and so the Evaluation Theorem says that:

$$\int_0^1 x\,dx = \frac{1}{2}x^2\Big|_0^1.$$

Again, the notation on the right means "plug 1 into the function $\frac{1}{2}x^2$, plug 0 into it, and subtract", so that we get

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

as the value of the integral in question. Note that this agrees with the value we found last time via geometry after interpreting this integral as the area of a triangle.

The same antiderivative also works when evaluating $\int_1^2 x\,dx$, which we also worked out last time using geometry. The Evaluation Theorem gives

$$\int_1^2 x\,dx = \frac{1}{2}x^2\Big|_1^2 = \frac{1}{2}(2)^2 - \frac{1}{2}(1)^2 = \frac{3}{2},$$

agreeing with the value we saw last time.

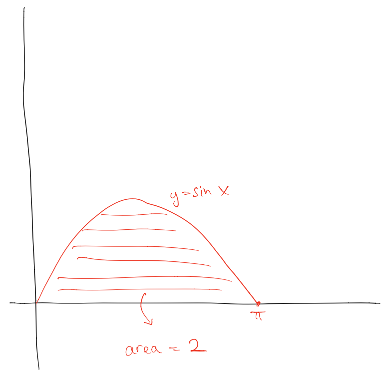
Example 2. We evaluate

$$\int_0^\pi \sin x\, dx.$$

Again, all we need is a function whose derivative is $\sin x$. We know that the derivative of $\cos x$ gives $-\sin x$, so this almost works. To get what we want we simply multiply by an extra negative, so we see that $-\cos x$ is a function whose derivative is $\sin x$. The Evaluation Theorem gives

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

(Be careful keeping track of the negatives!) Geometrically, this gives the area of the following region under the graph of $f(x) = \sin x$:



which is an area which is pretty much impossible to determine without the use of an integral and the Evaluation Theorem.

If instead we had

$$\int_0^\pi 5\sin x\, dx,$$

we would need a function whose derivative is $5\sin x$. Above we saw that $-\cos x$ has derivative $\sin x$, so $-5\cos x$ will have derivative $5\sin x$. This illustrates a general fact, namely that

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

when c is a constant. That is, constants can be "pulled out" of integrals, similarly to how constants can be pulled out of derivatives. Thus in this case we get:

$$\int_0^\pi 5\sin x\, dx = 5\int_0^\pi \sin x\, dx = -5\cos x\Big|_0^\pi = (-5\cos\pi) - (-5\cos 0) = 10.$$

Example 3. Finally we determine the value of

$$\int_0^\pi (5\sin x + x^2)\, dx.$$

So, we need an antiderivative of $5\sin x + x^2$. The key point is that we already know that $-5\cos x$ is an antiderivative of $5\sin x$, so if we also had an antiderivative of x^2, adding it to $-5\cos x$ would give an antiderivative of $5\sin x + x^2$. This comes from the fact that

$$\text{derivative of } (f+g) = (\text{derivative of } f) + (\text{derivative of } g).$$

In terms of integrals, this corresponds to the fact that

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

meaning that integrals of sums can always be split up into pieces. In our case, this means that

$$\int_0^\pi (5\sin x + x^2)\,dx = \int_0^\pi 5\sin x\,dx + \int_0^\pi x^2\,dx.$$

The first integral was computed to be 10 in Example 2. For the second we need a function whose derivative is x^2; x^3 almost works except that this gives us a 3 in front, so we throw in a $\frac{1}{3}$ in front to balance this out. Thus:

$$\begin{aligned}\int_0^\pi (5\sin x + x^2)\,dx &= \int_0^\pi 5\sin x\,dx + \int_0^\pi x^2\,dx \\ &= -5\cos x\Big|_0^\pi + \frac{1}{3}x^3\Big|_0^\pi \\ &= 10 + \left(\frac{1}{3}\pi^3 - \frac{1}{3}0^3\right) \\ &= 10 + \frac{1}{3}\pi^3.\end{aligned}$$

Indefinite integrals. Since computing integrals comes down to finding antiderivatives, it would be enormously helpful to have the basic antiderivatives engrained in our minds. These can be found in Table 1 of Section 5.3, and are all ones which you eventually want to have memorized. We'll look at more examples next time.

As a matter of notation, because of the connection between integrals and antiderivatives, we use

$$\int f(x)\,dx$$

to refer to the possible antiderivatives of the function f. This is what is known as an *indefinite integral*, to distinguish it from the definite integrals we've seen up to now. The difference is that a definite integral

$$\int_a^b f(x)\,dx$$

gives a number as a result, while an indefinite integral

$$\int f(x)\,dx \quad \text{(with no bounds on the integral sign)}$$

gives a collection of functions as a result, namely the functions which have f as their derivatives. You'll see in the book that all of these basic indefinite integral formulas include a "$+C$" at the end, which comes from the fact that adding a constant to a function does not alter its derivative. Later on we'll see situations where this "$+C$" plays an important role.

Lecture 4: The Fundamental Theorem of Calculus

Warm-Up 1. We evaluate the integral

$$\int_1^3 (2e^x - 2x^6 + \sec^2 x)\,dx.$$

To do so we need to find an antiderivative of the *integrand* $2e^x - 2x^6 + \sec^2 x$. (The term *integrand* just refers to the function which is being integrated.) First, $2e^x$ is its own derivative:

$$(2e^x)' = 2e^x.$$

Now, to get the x^6 in the second term we need to take the derivative of x^7:

$$(x^7)' = 7x^6.$$

To get rid of the 7 in front we divide by 7 throughout:

$$\left(\frac{1}{7}x^7\right)' = x^6,$$

and finally to get $-2x^6$ as required we multiply through by -2:

$$\left(-\frac{2}{7}x^7\right) = -2x^6.$$

Finally, the derivative of $\tan x$ is $\sec^2 x$:

$$(\tan x)' = \sec^2 x.$$

Thus $2e^x - \frac{2}{7}x^7 + \tan x$ is an antiderivative of $2e^x - 2x^6 + \sec^2 x$. Since adding any constant to $2e^x - \frac{2}{7}x^7 + \tan x$ still gives the same derivative (since the derivative of a constant is zero), we see that the most general antiderivative of $2e^x - 2x^6 + \sec^2 x$ is

$$\int (2e^x - 2x^6 + \sec^2 x)\, dx = 2e^x - \frac{2}{7}x^7 + \tan x + C$$

where C denotes an arbitrary constant. Recall that this integral notation without any bounds refers to the *indefinite integral* of $2e^x - 2x^6 + \sec^2 x$, which, by definition, means the most general antiderivative of $2e^x - 2x^6 + \sec^2 x$. (Indefinite integrals always include this $+C$ term, and later we will see instances where this $+C$ actually matters.)

To evaluate the original integral, all we need is *an* antiderivative of $2e^x - 2x^6 + \sec^2 x$. Thus we can just take the constant C to be zero; in other words, $2e^x - \frac{2}{7}x^7 + \tan x$ is one possible antiderivative of $2e^x - 2x^6 + \sec^2 x$, which is all we need in order to evaluate a definite integral. Thus:

$$\begin{aligned}\int_1^3 (2e^x - 2x^6 + \sec^2 x)\, dx &= \left(2e^x - \frac{2}{7}x^7 + \tan x\right)\Bigg|_1^3 \\ &= \left(2e^3 - \frac{2}{7}\cdot 3^7 + \tan 3\right) - \left(2e - \frac{2}{7} + \tan 1\right).\end{aligned}$$

Note that if had used, say, $2e^x - \frac{2}{7}x^7 + \tan x + 3$ as an antiderivative of $2e^x - 2x^6 + \sec^2 x$ instead of the one we used, the extra $+3$ term would have disappeared anyway since it would show up once when we plug in the upper bound of 3 but then again when we plug in the lower bound of 1, so that after subtracting the $+3$ would be no more. This is why the $+C$ term in an indefinite integral won't matter when computing a definite integral.

Warm-Up 2. Suppose that water is flowing in/out of a tank at a rate of $f(t) = 3t^3 - \frac{1}{t^2}$ cubic meters per second. We want to determine the total amount of water which flowed in/out of the tank over the time period of $t = 1$ second to $t = 4$ seconds, or in other words the *net change* in the volume over this time period. The point is that this is simply asking to compute the integral

$$\int_1^4 \left(3t^3 - \frac{1}{t^2}\right) dt.$$

Indeed, we've alluded to the idea before that "net change" questions can also be phrased in terms of integrals (such as the type of distance question we saw in the first lecture), so that integrals are not only meant to have applications to computing area. (Note that the version I first gave in class had 0 as the lower bound of the integral, which is nonsense since the $\frac{1}{t^2}$ portion of the function we are integrating isn't even defined at 0. This is why I changed the lower bound to 1 instead.)

We can now see why this is true from the Evaluation Theorem. Write the Evaluation Theorem as

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

The right side is the net change in F from a to b. The $F'(x)$ on the left measures the rate of change of F, so this equation says that integrating the rate of change precisely results in the net change. In our specific problem, $f(t) = 3t^3 - \frac{1}{t^2}$ is the rate of change of the volume function $F(t)$, so integrating it from 1 to 4 should give the net change in F itself.

Note also something we mentioned last time: the fact that we are using t as the variable makes no difference to the procedure for evaluating this integral, and the value of this integral is the same as the one for

$$\int_1^4 \left(3x^3 - \frac{1}{x^2}\right) dx, \text{ or } \int_1^4 \left(3☺^3 - \frac{1}{☺^2}\right) d☺.$$

We use t in this case simply because we are interpreting this variable as "time".

So, we compute the given integral. We have:

$$(t^4)' = 4t^3, \text{ so } \left(\frac{3}{4}t^4\right)' = 3t^3.$$

Now, to find an antiderivative for $\frac{1}{t^2}$, we write this function as t^{-2} instead. Then the same differentiation rules apply: to get a power of -2 we need to start with a power of -1:

$$(t^{-1})' = -t^{-2}.$$

Thus, $\frac{1}{t}$ is an antiderivative for $-\frac{1}{t^2}$, so all together

$$\frac{3}{4}t^4 + \frac{1}{t} \text{ is an antiderivative for } 3t^3 - \frac{1}{t^2}.$$

The most general antiderivative, if we had been asked for it, is

$$\int \left(3t^3 - \frac{1}{t^2}\right) dt = \frac{3}{4}t^4 + \frac{1}{t} + C.$$

To finish computing our specific integral, we have:

$$\begin{aligned}\int_1^4 \left(3t^3 - \frac{1}{t^2}\right) dt &= \left(\frac{3}{4}t^4 + \frac{1}{t}\right)\Bigg|_1^4 \\ &= \left(\frac{3}{4}\cdot 4^4 + \frac{1}{4}\right) - \left(\frac{3}{4} + 1\right)\end{aligned}$$

$$= \frac{770}{4}.$$

Thus the net change in volume of water in the tank between 1 and 4 seconds was $\frac{770}{4}$ cubic meters.

Fundamental Theorem of Calculus. Recall the idea of an *area function* from last time: for a function f, we can define a new function F by setting the value of F at x to be:

$$F(x) = \int_a^x f(t)\,dt.$$

This value gives the net area between the graph of f and the horizontal axis from a fixed point a up to some variable point x. As x changes, this net area changes as well. The *Fundamental Theorem of Calculus* (or Part I of this theorem, to distinguish it from Part II, which is also known as the Evaluation Theorem), tells us what happens when we take the derivative of this area function. The result is that

$$\text{for } F(x) = \int_a^x f(t)\,dt, \text{ we have } F'(x) = f(x).$$

Thus, differentiating the function obtained by integrating another function gives the function which was originally being integrated.

Now, at first glance this doesn't seem to be saying much, since we've already said before in the Evaluation Theorem that integrating a function should give its antiderivative, and of course it's true that differentiating this antiderivative gives the original function. The point, which we'll elaborate on next time, is that the *reason* why the Evaluation Theorem works and the reason why integrating a function amounts to finding an antiderivative is *because* of the Fundamental Theorem of Calculus. This specific version is called Part I since it logically is supposed to come before Part II. Again, we'll elaborate on this next time.

Integration and differentiation do opposite things. Before looking at some examples of how to use the Fundamental Theorem of Calculus, we emphasize what it and the Evaluation Theorem really say. The Evaluation Theorem says

$$\int_a^b F'(x)\,dx = F(b) - F(a)$$

and the Fundamental Theorem says

$$\frac{d}{dx}\left(\int_a^x f(t)\,dt\right) = f(x).$$

You should read the first as saying the following: start with a function F, take the derivative, and then take the integral, and you get back the original F itself. The second equation above says: start with a function f, take the integral, and then take the derivative, and you get back the original f itself. The point is that these together say that integration and differentiation are "inverse" operations, meaning that one does the opposite of what the other one does.

Example 1. Define a function F by

$$F(x) = \int_0^x t^3\,dt.$$

The Fundamental Theorem of Calculus says that $F'(x) = x^3$. Note that in this case f is the function $f(t) = t^3$, and the final conclusion that $F'(x) = f(x)$, where $f(x)$ means that we evaluate f, not at t, but at the upper bound x of the integral in question. Saying that

$$F'(x) = t^3$$

is incorrect; we must *always* evaluate the result at the upper bound of the integral since this is what the Fundamental Theorem of Calculus actually says.

Example 2. Define a function G by

$$G(x) = \int_0^x e^{t^2}\,dt.$$

In the previous example we can actually compute $\int_0^x t^3\,dt$ directly by finding an antiderivative of t^3, but the point is that in this case this is not possible: it is *not* possible to write down an explicit antiderivative of e^{t^2}. (Surprising, no?) Note that here we are not saying that no one has ever found a way to write down an explicit antiderivative, but rather that is has been *proven* that it is not possible to do so.

So, this leaves us asking: how do we even know that e^{t^2} has an antiderivative? The reason we know this is true is that the Fundamental Theorem of Calculus tells us that the function G defined above *is* an antiderivative of e^{x^2} since the theorem says that

$$G'(x) = e^{x^2}.$$

Hopefully this sheds some light on why the Fundamental Theorem is even worth stating at all: there are many examples of functions for which an explicit antiderivative cannot be written down, and in such cases the only type of expression we have for such an antiderivative is the one given by the Fundamental Theorem of Calculus itself.

Example 3. Define H by

$$H(x) = \int_1^{x^2} t^2\,dt.$$

We want to compute $H'(x)$. Is the answer obtained simply by evaluating the integrand t^2 as $t = x^2$, giving $H'(x) = x^4$? Not quite! The point here is that the Fundamental Theorem of Calculus only applies to integrals of a specific form: integrating from a constant as a lower bound to the variable x itself as the upper bound, whereas here we have an upper bound of x^2. We have to be careful with such an expression.

To see what to do, let's actually consider the function

$$F(u) = \int_1^u t^2\,dt.$$

I'm using u here to denote the variable instead of x so that we don't get mixed up with $H(x)$, which we'll come back to in a second. The Fundamental Theorem of Calculus says that

$$F'(u) = u^2,$$

which applies since the upper bound in the integral defining F is just u and not something like u^2. The relation between F and the function H we want is that H is what we get when we evaluate F at $u = x^2$:

$$H(x) = F(x^2) = \int_1^{x^2} t^2\,dt.$$

So, what we are really trying to do is compute the derivative of $H(x) = F(x^2)$. This is precisely the type of derivative to which the chain rule applies!

Indeed, thinking of F as the "outer" function and x^2 as the "inner" function, the chain rule says that:

$$H'(x) = \underbrace{F'(x^2)}_{\text{derivative of outer}} (\underbrace{\text{derivative of } x^2}_{\text{derivative of inner}}).$$

To be clear, the "derivative of outer" really means that derivative of $F(u)$ evaluated at $u = x^2$. Since $F'(u) = u^2$ by the Fundamental Theorem of Calculus, $F'(x^2) = (x^2)^2 - x^4$ when we set $u = x^2$. The derivative of x^2 is $2x$, so all together we get

$$H'(x) = x^4(2x) = 2x^5$$

as the derivative of

$$H(x) = \int_1^{x^2} t^2\,dt.$$

Again, the point is that in this case the upper bound is not simply x, so computing such a derivative requires both the Fundamental Theorem of Calculus *and* the chain rule.

Example 4. Define

$$G(x) = \int_{-1}^{\sin x} \cos(t^2)\,dt.$$

We want to compute $G'(x)$, and again the answer won't be simply what you get when you evaluate $\cos(t^2)$ at $t = \sin x$; we have to take into consideration the fact that the upper bound on the integral is $\sin x$ and not just x. Note that $\cos(t^2)$ is also a function for which it is not possible to write down an explicit antiderivative, so that in order to work with this integral we absolutely need the Fundamental Theorem of Calculus Part I since the Evaluation Theorem does not apply.

Consider the function

$$F(u) = \int_{-1}^{u} \cos(t^2)\,dt.$$

The Fundamental Theorem gives

$$F'(u) = \cos(u^2).$$

The relation between F and G is that G is what you get when you plug in $u = \sin x$ into F:

$$G(x) = F(\sin x).$$

The chain rule gives

$$G'(x) = F'(\sin x)(\text{derivative of } \sin x).$$

The first factor is $F'(u)$ evaluated at $u = \sin x$, and the second is $\cos x$, so overall we get

$$G'(x) = \cos(\sin^2 x)\cos x$$

as the derivative of G.

Example 5. Finally, we find the derivative of

$$F(x) = \int_{2x}^{x^2} e^t\,dt.$$

There are a few things to note here. First, it could be that depending on x, the lower bound $2x$ is actually larger than the upper bound x^2. (For instance, this happens when $x = 1$.) So, the bounds on this integral seem to be in the "wrong" order. But this is simple to deal with: the notation

$$\int_a^b f(x)\,dx \text{ when } b < a \textit{ means} \text{ by definition } -\int_b^a f(x)\,dx.$$

In other words, to compute an integral when the bounds are in the "wrong" order we simply switch the bounds around and change the sign of the integral. So, the expression for $F(x)$ makes sense even when $2x > x^2$.

More importantly, this integral is not the type to which the Fundamental Theorem of Calculus applies, not only because of the x^2 upper bound but also because the lower bound is not constant. (Look back to all examples we've said so far and note that the lower bounds were all constant, which is necessary in order to be able to apply the Fundamental Theorem.) So, the idea is that we must rewrite the expression for F in way which involves integrals with constant lower bounds. But this is also simple to do: we simply introduce some intermediate constant, say 0, and break up the interval of integration:

$$\int_{2x}^{x^2} e^t\,dt = \int_{2x}^{0} e^t\,dt + \int_0^{x^2} e^t\,dt.$$

Now, you may argue that 0 might not be between $2x$ and x^2, so that the interval $[2x, x^2]$ is not made up of the interval $[2x, 0]$ together with the interval $[0, x^2]$. This is all true, but is not a problem: in a situation such as one where

$$0 < 2x < x^2,$$

we would have

$$\int_0^{x^2} = \int_0^{2x} + \int_{2x}^{x^2},$$

and rearranging gives

$$\int_0^{x^2} - \int_0^{2x} = \int_{2x}^{x^2}, \text{ which is the same as } \int_0^{x^2} + \int_{2x}^{0} = \int_{2x}^{x^2}.$$

Thus, even when $0 < 2x < x^2$, it would still be true that

$$\int_{2x}^{x^2} e^t\,dt = \int_{2x}^{0} e^t\,dt + \int_0^{x^2} e^t\,dt.$$

So, we never have to worry about whether or not the "intermediate" term we're using with which to split up an integral is actually "between" our given bounds: splitting up still works. (This is essentially the reason why we set $\int_a^b$ to be $-\int_b^a$ when $a > b$.)

Thus the function F can be written as

$$F(x) = \int_{2x}^{0} e^t\,dt + \int_0^{x^2} e^t\,dt = -\int_0^{2x} e^t\,dt + \int_0^{x^2} e^t\,dt,$$

and *now* we have the type of expressions to which the Fundamental Theorem of Calculus applies. Together with the chain rule, we thus get:

$$\begin{aligned} F'(x) &= -e^{2x}(\text{derivative of } 2x) + e^{x^2}(\text{derivative of } x^2) \\ &= -2e^{2x} + 2xe^{x^2}. \end{aligned}$$

Lecture 5: Computing Integrals via Substitution

Warm-Up. We compute the derivative (with respect to x) of the function F defined by

$$F(x) = \int_{x^2}^{\sin x} te^t\,dt.$$

(This is different than the specific examples we looked at in class, but it still gets the same idea across.) We would like to use the Fundamental Theorem of Calculus, but the given function is not in the form to which that theorem applies just yet. Our first goal is rewrite the expression defining F so that it only involves terms which look like

$$\int_{\text{constant}}^{\text{something depending on } x}$$

since these are the types of things to which the Fundamental Theorem is applicable.

First we can introduce an intermediate constant, say 2:

$$\int_{x^2}^{\sin x} = \int_{x^2}^{2} + \int_{2}^{\sin x}.$$

(Note that here I'm focusing on the interval being integrated over, so for now I'm ignoring the function being integrated in the notation.) There was nothing special about 2, and we could have instead used

$$\int_{x^2}^{\sin x} = \int_{x^2}^{10} + \int_{10}^{\sin x}.$$

The point is that this intermediate constant isn't going to matter anyway when we actually take the derivative. Also note that it is not necessarily true that 2 sits between the values of x^2 and $\sin x$:

$$x^2 < 2 < \sin x,$$

and it could be that (depending on x) we instead have something like:

$$2 < \sin x < x^2.$$

We saw last time that this wasn't an issue: in this case we could break things up as

$$\int_{2}^{x^2} = \int_{2}^{\sin x} + \int_{\sin x}^{x^2},$$

and then rearrange things to get

$$\int_{x^2}^{\sin x} = \int_{x^2}^{2} + \int_{2}^{\sin x}$$

as we wanted. Finally, switching the bounds changes the sign, so

$$\int_{x^2}^{\sin x} = -\int_{2}^{x^2} + \int_{2}^{\sin x},$$

so our defining expression for F can instead be written as

$$F(x) = -\int_{2}^{x^2} te^t\,dt + \int_{2}^{\sin x} te^t\,dt.$$

Now we can differentiate each of these portions, using the fact that

$$\left(\int_{\text{constant}}^{\text{g(x)}} \text{integrand}\right)' = \text{integrand}(g(x)) \text{ times } g'(x);$$

that is, the derivative is obtained by evaluating the function being integrated at the upper bound $g(x)$ and multiplying by the derivative of $g(x)$ itself. Recall from the examples we saw last time that this comes from the chain rule, which is needed since the upper bound is not simply x but rather some function depending on x. **Ask if this is still unclear!** Thus, applying this fact to each piece of

$$F(x) = -\int_2^{x^2} te^t\,dt + \int_2^{\sin x} te^t\,dt$$

gives

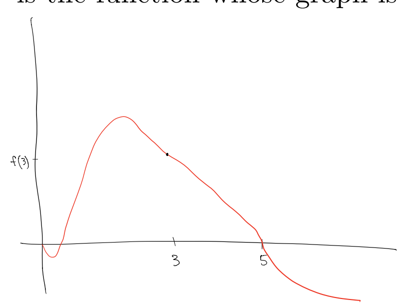
$$F'(x) = -(x^2 e^{x^2})(2x) + [(\sin x)e^{\sin x}](\cos x) = -2x^3 e^{x^2} + e^{\sin x}\sin x \cos x.$$

Why the Fundamental Theorem is fundamental. We now elaborate on something mentioned in passing last time. To recall, the Evaluation Theorem says that to compute an integral we should first find an antiderivative. We have said nothing about why this works until now, but the point is that it works *because* of the Fundamental Theorem of Calculus! Indeed, the Fundamental Theorem precisely says the function obtained by integrating f *is* a function whose derivative is f itself, so that this integrated function *is* an antiderivative of f. Hence, in order to compute an integral, we should find this antiderivative via whatever methods we have available. This is why the Evaluation Theorem is "Part II" of the Fundamental Theorem of Calculus: Part II is a consequence of Part I, which is really the key idea.

Of course, we won't give a proof of the Fundamental Theorem of Calculus (Part I) in this course, but nonetheless here is an example which I hope illustrates why this theorem is "plausible"; that is, why is it that differentiating

$$F(x) = \int_a^x f(t)\,dt$$

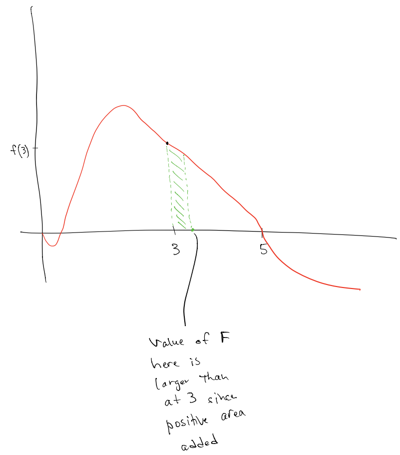
should just give $f(x)$? Suppose f is the function whose graph is drawn below:



Recall that the function F defined above measures the net area between the graph of f and the x-axis from a up to some varying x. Now, for the function f drawn above, the Fundamental Theorem would say that

$$F'(3) = f(3) \text{ is positive.}$$

Thinking back to what we know about derivatives, this should conceivably mean that F is increasing at $x = 3$, and indeed I claim this makes sense based on the graph of f! The value of $F(3)$ measures the net area up until 3, and the point is that if we then go a bit beyond $x = 3$ to the right, we will be adding a *positive* area to this net area:

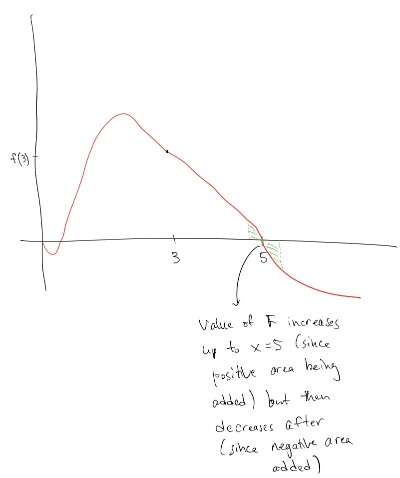


Thus, the net area up to a point a bit to the right of 3 will be larger than the net area up to 3, so that F evaluated at this new point will give a value larger than $F(3)$. Hence, F does increase through $x = 3$.

Similarly, the Fundamental Theorem would say that

$$F'(5) = f(5) = 0.$$

From what we know about derivatives, this suggests that F might have either a maximum or minimum at $x = 5$, and again I claim that we can tell it has a maximum based on the picture above. Right up to $x = 5$ the value of $F(x)$ increases since up to this point we are adding positive areas; however, once we move beyond 5 we will be adding negative areas, which will cause the value of F to decrease:



Thus, F increases up to $x = 5$, then begins to decrease after, which is precisely what should happen when F has a maximum at $x = 5$

The point is that the fact that $F'(x) = f(x)$ is something we can make sense of based on how the net area function F is changing; F is increasing when f is positive, F is decreasing when f is negative, and F has a maximum or minimum when f is zero, which all together says that f does seem to behave in a manner similar to how the derivative of F should behave.

Substitution. We now return to the problem of finding antiderivatives, which we'll spend a good week or two on. The first method we'll consider is what's called *substitution*, or sometimes u-*substitution* based on the fact that we normally use u to denote the substitution we are making.

The goal of substitution is to transform a given integral into a new one in terms of some new variable, with the hope of having this new integral be simpler to compute.

Consider for instance the integral

$$\int xe^{x^2}\,dx.$$

You can check by hand that $\frac{1}{2}e^{x^2}$ has derivative equal to xe^{x^2}, so $\frac{1}{2}e^{x^2}+C$ is the general antiderivative of xe^{x^2}:

$$\int xe^{x^2}\,dx = \frac{1}{2}e^{x^2} + C.$$

The point is that we want to have a way to figure this out without having to resort to any guesswork; namely, how exactly can we find that $\frac{1}{2}e^{x^2}$ is the correct thing to consider when starting with xe^{x^2}?

Here is the process. We introduce a new variable u by setting

$$u = x^2.$$

Note that with this the e^{x^2} portion of our integral becomes simply e^u. Our goal is now to rewrite the given integral in terms of u instead. Recall from what we know about the derivatives that the *differential* of u is given by:

$$du = 2x\,dx,$$

which is obtained by differentiating both sides of $u = x^2$. After dividing both sides by 2 we get

$$\frac{1}{2}\,du = x\,dx.$$

The point is that the $x\,dx$ on the right makes up the rest of the integral we are considering: the e^{x^2} term is e^u, and the remaining $x\,dx$ becomes $\frac{1}{2}\,du$. Thus we can rewrite our given integral as

$$\int xe^{x^2}\,dx = \int \frac{1}{2}e^u\,du.$$

This integral on the right is now straightforward to compute, where the key point is that we are integrating with respect to the same variable u which our function is written in terms of; in other words, the integral $\int e^u\,du$ is the same as that of $\int e^x\,dx$ only that we denote the variable by u instead of x. We have

$$\int e^u\,du = e^u + C,$$

so all together we get

$$\int xe^{x^2}\,dx = \int \frac{1}{2}e^u\,du = \frac{1}{2}e^u + C.$$

As a final step we can express the result back in terms of x by substituting back in for u: we had originally set $u = x^2$, so

$$\int xe^{x^2}\,dx = \frac{1}{2}e^u + C = \frac{1}{2}e^{x^2} + C,$$

agreeing with the same we gave at the start without any motivation.

Why substitution works. As a general mantra, whenever we have some differentiation technique, integrating it will give rise to some analogous integration technique. In this setting, substitution is meant to be the analog of the chain rule. The key thing to is to recognize (if possible) the integral we are wanting to compute as being of the form

$$\int f'(g(x))g'(x)\,dx,$$

where a certain portion $g'(x)$ looks to be the derivative of some other portion $g(x)$. (Note in the xe^{x^2} example above, the x in front is similar to the derivative of the x^2 portion.) This type of expression shows up in the chain rule:

$$f(g(x))' = f'(g(x))g'(x),$$

which then gives

$$\int f'(g(x))g'(x)\,dx = \int f(g(x))'\,dx = f(g(x)) + C$$

after integrating.

To phrase this in terms of u, set

$$u = g(x).$$

Then $du = g'(x)\,dx$, so the expression

$$\int f'(g(x))g'(x)\,dx \quad \text{is the same as} \quad \int f'(u)\,du,$$

which works out to be $f(u)+C = f(g(x))+C$. The point is that after writing our original integral in terms of u, we can then simply integrate what we're left with respect to u (forgetting at this point about what u actually stood for), and then plug back in for u in the end. (Later after we talk a bit more about what integrals "really" mean, we'll come back and talk about what $du = g'(x)\,dx$ actually means.) Introducing u is meant to help us simplify our expression by isolating which part is the actually thing we have to integrate and which part comes from taking the derivative of something else.

Another example. We compute

$$\int 2x^3 \cos(5x^4)\,dx.$$

The first is to recognize part of this as coming from the derivative of another part; in this case, x^3 comes from the derivative of $5x^4$, so we set

$$u = 5x^4.$$

(Note that setting $u = x^4$ would also work, but in general to save on the amount of work we want u to take up as much as possible of our original expression; $u = 5x^4$ leads to $\cos(u)$ where as $u = x^4$ leads to $\cos(5u)$, so $u = 5x^4$ takes up more of the $\cos(5x^4)$ expression.) We get

$$du = 20x^3\,dx.$$

The $\cos(5x^4)$ part becomes $\cos u$, so we are left with the $2x^3\,dx$ part. Dividing both sides of

$$du = 20x^3\,dx$$

by 10 gives

$$\frac{1}{10}\,du = 2x^3\,dx,$$

which is precisely what we want. Thus

$$\int 2x^3 \cos(5x^4)\, dx = \int \frac{1}{10} \cos u \, du.$$

This final integral (with respect to u) is $\frac{1}{10} \sin u + C$, so after plugging back in for u we get

$$\int 2x^3 \cos(5x^4)\, dx = \frac{1}{10} \sin(5x^4) + C.$$

You can (and should!) verify that taking the derivative of $\frac{1}{10} \sin(5x^4)$ indeed gives $2x^3 \cos(5x^4)$.

Yet another example. We compute

$$\int_e^{e^2} \frac{1}{x \ln x}\, dx.$$

Forget the bounds for now. To use substitution we again recognize that part of our integrand comes from the derivative of another part; in this case, $\frac{1}{x}$ comes from the derivative of the $\ln x$ in the denominator. Thus we set

$$u = \ln x, \text{ in which case } du = \frac{1}{x}\, dx.$$

Thinking of our integrand as $\frac{1}{x} \frac{1}{\ln x}\, dx$, we thus have $\frac{1}{u}\, du$. Hence:

$$\int \frac{1}{x \ln x}\, dx = \int \frac{1}{u}\, du = \ln|u| + C = \ln|\ln x| + C.$$

Again, you should verify that the derivative of $\ln|\ln x|$ is $\frac{1}{x \ln x}$. Thus now taking the bounds into account, we get:

$$\int_e^{e^2} \frac{1}{x \ln x}\, dx = (\ln|\ln x|)\Big|_e^{e^2} = \ln(\ln(e^2)) - \ln(\ln e) = \ln 2 - \ln 1 = \ln 2.$$

Note that above we first worked out the antiderivative, expressed everything back in terms of x, and then used the bounds. Alternatively, once we get to

$$\int \frac{1}{x \ln x}\, dx = \int \frac{1}{u}\, du$$

we can work out new bounds on the u-integral in terms of what happens to u as x varies. In our case, the substitution we used was $u = \ln x$: when $x = e$ in the original integral, u will have the value $u = \ln e = 1$, and when $x = e^2$ in the original integral, u will have the value $u = \ln(e^2) = 2$. In other words, as x ranges from e to e^2, $u = \ln x$ will in turn range from 1 to 2. This gives the bounds we want on the u-integral:

$$\int_e^{e^2} \frac{1}{x \ln x}\, dx = \int_1^2 \frac{1}{u}\, du.$$

Computing the integral on the right gives

$$\int_1^2 \frac{1}{u}\, du = \ln|u|\Big|_1^2 = \ln 2 - \ln 1 = \ln 2,$$

which agrees with what we had before. This technique will always work: we can either compute definite integral by expressing everything in terms of x and using the original bounds of integration, or we can keep everything in terms of u and work out the new bounds on u by seeing what happens to u when x is evaluated at the original bounds.

Final example. Finally, we compute

$$\int x\sqrt{2+x}\,dx.$$

The new thing here is that there is no part of this which looks like the derivative of another part, since the x in front does not come from the derivative of the $2+x$ term under the square root. Nonetheless, let us see what happens if we use the substitution

$$u = 2 + x$$

anyway. Then $du = dx$, so the $\sqrt{2+x}$ and dx portions of the original integral are taken care of: they becomes $\sqrt{u}$ and du respectively. But what about the remaining x in front? The point is that we do know how to express this x in terms of u: since $u = 2+x$, $x = u-2$. Thus we rewrite our entire integral, originally in terms of x, as:

$$\int x\sqrt{2+x}\,dx = \int (u-2)\sqrt{u}\,du.$$

Did this get us anywhere? Yes! Now using $\sqrt{u} = u^{1/2}$, we can write the resulting integral as

$$\int (u-2)u^{1/2}\,du = \int (u^{3/2} - 2u^{1/2})\,du,$$

which is now possible to integrate directly. Note the key difference between this expression and the original one: in the original we had

$$x(2+x)^{1/2},$$

and there is no simple way to distribute the x through inside the parenthesis, whereas now we have

$$(u-2)u^{1/2},$$

where distributing is possible. We end up with:

$$\int x\sqrt{2+x}\,dx = \int (u^{3/2} - 2u^{1/2})\,du = \frac{2}{5}u^{5/2} - \frac{4}{3}u^{3/2} + C = \frac{2}{5}(2+x)^{5/2} - \frac{4}{3}(2+x)^{3/2} + C.$$

As a check, the derivative of the expression on the right is

$$\frac{2}{5}\frac{5}{2}(2+x)^{3/2} - \frac{4}{3}\frac{3}{2}(2+x)^{1/2} = (2+x)(2+x)^{1/2} - 2(2+x)^{1/2} = (2+x-2)(2+x)^{1/2},$$

which is $x\sqrt{2+x}$ as expected. The moral is that, really, we can make any kind of substitution we like, as long as we can write our integral completely in terms of our new variable.

Lecture 6: Integration by Parts

Warm-Up 1. We compute

$$\int \frac{x\cos(x^2)}{\sqrt{3+\sin(x^2)}}\,dx.$$

Note that the entire numerator comes from the derivative of $\sin(x^2)$. Thus we make the substitution

$$u = 3+\sin(x^2).$$

(Recall: u should take up as much of the given expression as possible, so $u = 3+\sin(x^2)$ is better than $u = \sin(x^2)$.) Then

$$du = 2x\cos(x^2)\,dx,$$

so

$$\frac{1}{2}du = x\cos(x^2)\,dx.$$

Thus our original integral becomes

$$\int \frac{x\cos(x^2)}{\sqrt{3+\sin(x^2)}}\,dx = \int \frac{1}{2}\frac{1}{\sqrt{u}}\,du.$$

Writing $\frac{1}{\sqrt{u}}$ as $u^{-1/2}$, we have

$$\frac{1}{2}\int \frac{1}{\sqrt{u}}\,du = \frac{1}{2}\int u^{-1/2}\,du = u^{1/2}+C.$$

Thus

$$\int \frac{x\cos(x^2)}{\sqrt{3+\sin(x^2)}}\,dx = \sqrt{u}+C = \sqrt{3+\sin(x^2)}+C.$$

As a check, taking the derivative of the function we end up with gives

$$\frac{1}{2}(3+\sin(x^2))^{-1/2}(2x\cos(x^2)) = \frac{x\cos(x^2)}{\sqrt{3+\sin(x^2)}}$$

as expected.

Warm-Up 2. We compute

$$\int x^3\sqrt{x^2+1}\,dx.$$

This is one of those situations where the x^3 in front doesn't come directly from the derivative of the term under the square root, but where we set u equal to x^2+1 anyway and in the end write everything in terms of u. We have

$$u = x^2+1, \text{ so } du = 2x\,dx, \text{ and } \frac{1}{2}du = x\,dx.$$

We need to come up with $x^3\,dx$, whereas so far we only have $x\,dx$. Thus we must multiply through by x^2, and the point is that we do know what x^2 equals in terms of u: since $u = x^2+1$, $x^2 = u-1$. Thus

$$x^3\,dx = x^2(x\,dx) = (u-1)\frac{1}{2}\,du.$$

Thus our original integral becomes

$$\int x^3\sqrt{x^2+1}\,dx = \int \frac{1}{2}(u-1)\sqrt{u}\,du.$$

Computing gives:

$$\frac{1}{2}\int (u^{3/2}-u^{1/2})\,du = \frac{1}{2}\left(\frac{2}{5}u^{5/2}-\frac{2}{3}u^{3/2}\right)+C,$$

so after plugging back in $u = x^2+1$ we get:

$$\int x^3\sqrt{x^2+1}\,dx = \frac{1}{5}(x^2+1)^{5/2}-\frac{1}{3}(x^2+1)^{3/2}+C.$$

You can (and should!) verify that taking the derivative of the expression we ended up with does indeed give $x^3\sqrt{x^2+1}$.

Warm-Up 3. Suppose f is a continuous function which satisfies $\int_0^9 f(x)\,dx = 4$. We compute

$$\int_0^3 xf(x^2)\,dx.$$

The point here is that we don't have a specific f in mind, but that the same method we've been using is applicable nonetheless. To see what to do, imagine that we did have a specific f in mind, say $f(x) = \sin x$. Then $f(x^2)$ becomes $\sin(x^2)$, so we would be looking at the integral

$$\int_0^3 x\sin(x^2)\,dx.$$

If this were the case, we can see that we should in fact make the substitution $u = x^2$. The point is that this is still true even when we have a unspecific f; the x in front of $xf(x^2)$ comes from the derivative of x^2, so $u = x^2$ is the correct substitution to make.

With $u = x^2$, we get $du = 2x\,dx$, so

$$\frac{1}{2}\,du = x\,dx.$$

Then our integral becomes

$$\int_0^3 xf(x^2)\,dx = \frac{1}{2}\int_0^9 f(u)\,du,$$

where the new bounds come from evaluating $u = x^2$ as the old bounds $x = 0$ and $x = 3$. Lo and behold this resulting integral is precisely the one we are told has the value 4 at the beginning! Keep in mind that

$$\int_0^9 f(x)\,dx \text{ and } \int_0^9 f(u)\,du$$

denote the same thing, just using different notation for the variable of integration. Thus, in our case we get

$$\int_0^3 xf(x^2)\,dx = \frac{1}{2}\int_0^9 f(u)\,du = \frac{1}{2}(4) = 2$$

as the desired value.

Integration by parts. Suppose we want to compute

$$\int xe^x \, dx.$$

I claim that so far we don't have a good way of doing this. For instance, substitution may not work since the x in front is not the derivative of the exponent x in the e^x term. (Contrast this with $\int xe^{x^2} \, dx$, where substitution would work nicely.) Instead we could try setting $u = e^x$, in which case $du = e^x \, dx$. Since then $x = \ln u$, the given integral could be written as

$$\int \ln u \, du.$$

But again we're stuck now, as we don't have a good way of computing this integral yet either.

The integration technique we now look at is called *integration by parts*, and is meant to transform a given integral into a simpler to compute integral. This method is the integration-analog of the product rule, which says:

$$[f(x)g(x)]' = f(x)g'(x) + f'(x)g(x).$$

Rearranging gives

$$f(x)g'(x) - [f(x)g(x)]' - f'(x)g(x),$$

and taking antiderivatives of both sides gives:

$$\int f(x)g'(x) \, dx = f(x)g(x) - \int f'(x)g(x) \, dx.$$

The left hand side is meant to denote the integral we are originally given, so this says that if we can express the given integrand as

(function)(derivative of another function),

the formula above gives us a way to rewrite what we have.

To make notation simpler, it is common to introduce the variables

$$u = f(x) \text{ and } v = g(x),$$

so that $du = f'(x) \, dx$ and $dv = g'(x) \, dx$. The integration by parts formula then becomes

$$\int u \, dv = uv - \int v \, du.$$

The starting point in using this technique is to decide what plays the role of u and what plays the role of dv. Let's look at some examples.

Example 1. We return to the integral

$$\int xe^x \, dx.$$

Again, we want to think of xe^x as the product of some function u with the derivative dv of some other function v. Let's set

$$u = x \text{ and } dv = e^x \, dx.$$

Note that with these choices $u\,dv = xe^x\,dx$ is indeed the thing we are integrating in the given integral. Since $u = x$ we get $du = dx$. Now, v should be a function whose differential is $dv = e^x\,dx$, or in other words v should be an antiderivative of e^x; taking $v = e^x$ works. To summarize, we have:

$$\begin{aligned} u &= x & \qquad v &= e^x \\ du &= dx & \qquad dv &= e^x\,dx. \end{aligned}$$

Plugging this into the integration by parts formula

$$\int u\,dv = uv - \int v\,du$$

gives

$$\int xe^x\,dx = xe^x - \int e^x\,dx.$$

Thus, as claimed, integration by parts has rewritten our original integral in a different form. The point now is that the remaining integral $\int e^x\,dx$ is simply to compute: it is $e^x + C$. Thus all together we get

$$\int xe^x\,dx = xe^x - \int e^x\,dx = xe^x - e^x + C$$

as our answer. As a check, the derivative of the resulting expression is:

$$(xe^x - e^x + C)' = xe^x + e^x - e^x = xe^x,$$

where for the first term we use the product rule. Thus $xe^x - e^x + C$ is indeed the general antiderivative of xe^x.

Now, why did we set $u = x$ and $dv = e^x\,dx$ at the beginning, instead of say

$$u = e^x \text{ and } dv = x\,dx.$$

After all, these choices still give $u\,dv$ as $xe^x\,dx$ as we want. In this case, for du and v we get:

$$\begin{aligned} u &= e^x & \qquad v &= \frac{1}{2}x^2 \\ du &= e^x dx & \qquad dv &= x\,dx. \end{aligned}$$

Putting this into the integration by parts formula gives:

$$\int xe^x\,dx = \frac{1}{2}x^2e^x - \int \frac{1}{2}x^2e^x\,dx.$$

This is certainly a true statement, but the problem is that the remaining integral $\int x^2e^x\,dx$ is now not so simple to compute, and indeed it is more complicated than the original $\int xe^x\,dx$ we started with! The point is that we should aim to have the remaining $\int v\,du$ integral be simpler than the one we started with, which happens for $u = x, dv = e^x\,dx$ but not $u = e^x, dv = x\,dx$. The underlying reason here is that differentiating x gives some simpler, but anti-differentiating x gives something more complicated.

Example 2. We compute

$$\int x^2 \cos x\,dx.$$

We set

$$u = x^2 \text{ and } dv = \cos x\, dx.$$

Note that instead trying to use $u = \cos x, dv = x^2\, dx$ would lead to a more complicated second integral in the integration by parts formula, since integrating x^2 to get v will result in an x^3 term. With the choices above we get:

$$\begin{aligned} u &= x^2 & v &= \sin x \\ du &= 2x\, dx & dv &= \cos x\, dx \end{aligned}$$

so integration by parts gives

$$\int x^2 \cos x\, dx = x^2 \sin x - \int 2x \sin x\, dx.$$

Note that the remaining integral is indeed "simpler" than the original one in that it has a lower power of x.

How do we now compute this remaining integral? Via another application of integration by parts! Set

$$p = 2x \text{ and } dq = \sin x\, dx.$$

Note that I don't want to use u and dv here since this may cause confusion with the u and dv I used previously; with p and dq the integration by parts formula looks like:

$$\int p\, dq = pq - \int q\, dp.$$

We get:

$$\begin{aligned} p &= 2x & q &= -\cos x \\ dp &= 2\, dx & dq &= \sin x\, dx, \end{aligned}$$

so

$$\int 2x \sin x\, dx = -2x \cos x - \int -2 \cos x\, dx = -2x \cos x + 2 \int \cos x\, dx = -2x \cos x + 2 \sin x.$$

Putting this back into the $\int 2x \sin x\, dx$ portion of our original integration by parts application gives:

$$\begin{aligned} \int x^2 \cos x\, dx &= x^2 \sin x - \int 2x \sin x\, dx \\ &= x^2 \sin x - [-2x \cos x + 2 \sin x] + C \\ &= x^2 \sin x + 2x \cos x - 2 \sin x + C. \end{aligned}$$

Again, to be clear, the $\int 2x \sin x\, dx$ was itself computed via an integration by parts. To check that this is correct, we differentiate the result using some product rules:

$$(x^2 \sin x + 2x \cos x - 2 \sin x + C)' = 2x \sin x + x^2 \cos x + 2 \cos x - 2x \sin x - 2 \cos x = x^2 \cos x$$

as we would want to be the case.

Example 3. Finally we compute

$$\int \ln x\, dx.$$

Thinking of the integrand as $\ln x$ times $1\,dx$, we use

$$u = \ln x \text{ and } dv = dx.$$

Note that trying to use something like $dv = \ln x\,dx$ wouldn't work since in order to figure out v from this we would have to know the antiderivative of $\ln x$, which is precisely what it is we're trying to find in this problem. We get:

$$u = \ln x \qquad v = x$$
$$du = \frac{1}{x}\,dx \qquad dv = dx,$$

so

$$\int \ln x\,dx = x\ln x - \int x\frac{1}{x}\,dx = x\ln x - \int dx = x\ln x - x + C.$$

Note what happened here: the $v\,du$ portion ends up simplifying to give a simple integral to compute in the $\int v\,du$ portion of the integration by parts formula.

You can check that differentiating the resulting expression indeed gives $\ln x$, but recall actually that this is what you did on the second problem on Worksheet 1 from discussion. Now we know how to derive this antiderivative directly using integration by parts.

Lecture 7: Trigonometric Integrals

Warm-Up 1. We compute the definite integral

$$\int_1^2 2x^3 e^{x^2}\,dx$$

using integration by parts. First off, the bounds don't alter the methods since we can simply put bounds everywhere in the integration by parts formula:

$$\int_a^b u\,dv = uv\Big|_a^b - \int_a^b v\,du.$$

Now, we have to decide how we can express the expression we are integrating as one function times the derivative of another. We use:

$$u = x^2 \text{ and } dv = 2xe^{x^2}\,dx.$$

Note that this indeed gives the correct expression for $u\,dv$, and that the v term can be found by integrating

$$\int 2xe^{x^2}\,dx$$

using, say, a substitution; we get $v = e^{x^2}$. Thus we have:

$$u = x^2 \qquad v = e^{x^2}$$
$$du = 2x\,dx \qquad dv = 2xe^{x^2}\,dx.$$

Integration by parts gives:

$$\int_1^2 2x^3e^{x^2}\,dx = x^2e^{x^2}\Big|_1^2 - \int_1^2 2xe^{x^2}\,dx.$$

This remaining integral can be computed via a substitution, or by noting that it is precisely what we took to be dv, so that its integral is the v we had above. We get:

$$\begin{aligned}\int_1^2 2x^3e^{x^2}\,dx &= x^2e^{x^2}\Big|_1^2 - \int_1^2 2xe^{x^2}\,dx \\ &= x^2e^{x^2}\Big|_1^2 - e^{x^2}\Big|_1^2 \\ &= 4e^4 - e - (e^4 - e) = 3e^4\end{aligned}$$

as the final value.

Now, what else could we have tried? Going back to the original integral, say we had thought to try

$$u = 2x^3 \text{ and } dv = e^{x^2}\,dx.$$

The problem with this is that e^{x^2} is a function for which it is not possible to write down an explicit antiderivative! (Note that doesn't mean no one has yet a found a way to do it, but that it can in fact be *proven* that it not possible to do so.) So, we would be stuck when trying to use this as dv. This is at the heart of the reason why we set $dv = 2xe^{x^2}$, to get something for which v is actually possible to compute. The problem with using something like

$$u = e^{x^2} \text{ and } dv = 2x^3\,dx$$

is that the remaining integral after using integration by parts would actually be more complicated than the original one we're trying to compute! Indeed, this remaining integral would involve $x^5e^{x^2}$, which is worse than the original $x^3e^{x^2}$. The moral is that finding the correct u and dv to use isn't always straightforward, and is the kind of thing which becomes simpler with practice.

Nonetheless, here is another approach which also works. Let's first start with a regular substitution:

$$u = x^2 \text{ so } du = 2x\,dx.$$

Then

$$2x^3e^{x^2}\,dx = ue^u\,du,$$

so that our original integral becomes

$$\int_1^2 2x^3e^{x^2}\,dx = \int_1^4 ue^u\,du.$$

To compute the resulting integral in terms of u we can *now* do an integration by parts, say with

$$p = u \text{ and } dq = e^u\,du.$$

After working this out you end up with the same $3e^4$ value we found above. Thus in this method we still end up doing an integration by parts anyway, but we first use a substitution to rewrite the given integral as a simpler integral in terms of u.

Warm-Up 2. We compute

$$\int e^x \cos x\,dx.$$

This will illustrate one more type of technique which might show up when doing an integration by parts. First, we set $u = e^x, dv = \cos x\,dx$, so that:

$$u = e^x \qquad\qquad v = \sin x$$

$$du = e^x\,dx \qquad\qquad dv = \cos x\,dx.$$

Integration by parts gives:

$$\int e^x \cos x\,dx = e^x \sin x - \int e^x \sin x\,dx.$$

For this remaining integral we can again do an integration by parts:

$$p = e^x \qquad\qquad q = -\cos x$$
$$dp = e^x\,dx \qquad\qquad dq = \sin x\,dx.$$

We get:

$$\begin{aligned}\int e^x \cos x\,dx &= e^x \sin x - \int e^x \sin x\,dx \\ &= e^x \sin x - \left[-e^x \cos x - \int -e^x \cos x\,dx\right] \\ &= e^x \sin x + e^x \cos x - \int e^x \cos x\,dx.\end{aligned}$$

At this point it seem as if we're stuck, since if keep doing integration by parts we'll just keep flipping back and forth between $e^x \sin x$ integrals and $e^x \cos x$ integrals. However, here is the key observation: the remaining $\int e^x \cos x\,dx$ integral is precisely the integral on the left side we were originally trying to compute. Think of

$$\int e^x \cos x\,dx = e^x \sin x + e^x \cos x - \int e^x \cos x\,dx$$

as an equation where $\int e^x \cos x\,dx$ is the unknown quantity we want to find. In other words, if I think of $\int e^x \cos x\,dx$ as just some unknown variable:

$$☺ = \int e^x \cos x\,dx,$$

then the equation we have is

$$☺ = e^x \sin x + e^x \cos x - ☺.$$

We can solve this for ☺ by adding it to both sides to get:

$$2☺ = e^x \sin x + e^x \cos x.$$

In terms of what ☺ represents this says

$$2\int e^x \cos x\,dx = e^x \sin x + e^x \cos x,$$

so dividing by 2 gives the value we're looking for:

$$\int e^x \cos x\,dx = \frac{1}{2}(e^x \sin x + e^x \cos x) + C.$$

(The $+C$ was thrown on at the end as usual.) We can verify that this is the correct indefinite integral by taking the derivative of what we got:

$$\frac{1}{2}(e^x \sin x + e^x \cos x)' = \frac{1}{2}(e^x \sin x + e^x \cos x + e^x \cos x - e^x \sin x) = e^x \cos x.$$

Trigonometric Integrals. We now consider integrals which involve powers of the basic trig functions $\sin x, \cos x, \sec x, \tan x$. Such things are called *trigonometric integrals*, and the point here is that we can use well-known trig identities to transform such integrals into more easily computable forms. The key trig identities we use are:

$$\sin^2\theta + \cos^2\theta = 1,\ \sec^2\theta = \tan^2\theta + 1,\ \cos^2\theta = \frac{1}{2}(1+\cos 2\theta),\ \sin^2\theta = \frac{1}{2}(1-\cos 2\theta).$$

The first is one we all know and love, the second comes from dividing both sides of the first by $\cos^2\theta$, and the second and third are called the *half-angle* formulas. Note that these last two turn expressions involving sine or cosine squared into ones involving sine and cosine themselves with the angle doubled.

Example 1. We compute

$$\int \sin^3 x\,dx.$$

First note that normal substitution $u = \sin x$ doesn't work since this would give a $\cos x$ term in du, but there is no $\cos x$ term in the given integral which we would need to be able to introduce du. (If this was $\int \sin^3 x \cos x\,dx$ instead then a simple substitution $u = \sin x$ would indeed work nicely.)

The key here is that we can think of $\sin^3 x$ (which is $\sin x$ times itself three times) as

$$\sin^3 x = \sin^2 x \sin x,$$

and the $\sin^2 x$ term is something we can replace with something involving $\cos^2 x$, namely $1 - \cos^2 x$. We have:

$$\int \sin^3 x\,dx = \int \sin^2 x \sin x\,dx = \int (1-\cos^2 x)\sin x\,dx.$$

Now the resulting integral *is* one we can do using a substitution, say $u = \cos x$. With this we get $du = -\sin x\,dx$, so the given integral becomes

$$\begin{aligned}\int \sin^3 x\,dx &= \int (1-\cos^2 x)\sin x\,dx \\ &= \int -(1-u^2)\,du \\ &= -u + \frac{1}{3}u^3 + C \\ &= -\cos x + \frac{1}{3}\cos^3 x + C.\end{aligned}$$

To check that this is correct, we compute a derivative:

$$\left(-\cos x + \frac{1}{3}\cos^3 x\right)' = \sin x + \frac{3}{3}\cos^2 x(-\sin x) = \sin x(1-\cos^2 x) = \sin x \sin^2 x = \sin^3 x,$$

so that $-\cos x + \frac{1}{3}\cos^3 x + C$ is indeed the general antiderivative of $\sin^3 x$.

The goal here and for general trig integrals is to use a trig identity to rewrite the given integral in a different form, where one portion ends up being related to the derivative of another portion, after which a normal substitution should come in handy.

Example 2. We compute

$$\int \cos^2 x\,dx.$$

Again note that an ordinary substitution like $u = \cos x$ doesn't help since there is no $\sin x$ term as would be required by du. However, here we can use one of the half-angle formulas, namely

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

We get

$$\int \cos^2 x\, dx = \int \frac{1}{2}(1 + \cos 2x)\, dx = \frac{1}{2}\int (1 + \cos 2x)\, dx.$$

The 1 can be integrated easily, and the $\cos 2x$ can be integrated via a substitution $p = 2x$. We get

$$\int \cos 2x\, dx = \frac{1}{2}\sin 2x + C,$$

so

$$\int \cos^2 x\, dx = \frac{1}{2}\int (1 + \cos 2x)\, dx = \frac{1}{2}\left(x + \frac{1}{2}\sin 2x\right) + C = \frac{1}{2}x + \frac{1}{4}\sin 2x + C$$

is the desired indefinite integral. We can check this by computing a derivative:

$$\left(\frac{1}{2}x + \frac{1}{4}\sin 2x\right)' = \frac{1}{2} + \frac{2}{4}\cos 2x = \frac{1}{2}(1 + \cos 2x) = \cos^2 x$$

as expected.

Just to show a different use of a half-angle formula, note that we can use $\sin^2 x + \cos^2 x = 1$ to write the original integral as

$$\int \cos^2 x\, dx = \int (1 - \sin^2 x)\, dx = \int dx - \int \sin^2\, dx.$$

The $\sin^2 x$ can be computed using half-angle identity $\sin^2 = \frac{1}{2}(1 - \cos 2x)$. We get:

$$\begin{aligned}\int \cos^2 x\, dx &= \int dx - \int \sin^2\, dx \\ &= x - \frac{1}{2}\int (1 - \cos 2x)\, dx \\ &= x - \frac{1}{2}\left(x - \frac{1}{2}\sin 2x\right) + C \\ &= \frac{1}{2}x + \frac{1}{4}\sin 2x + C,\end{aligned}$$

which is the same indefinite integral we found before.

Lecture 8: Trigonometric Substitution

Warm-Up 1. We compute the integral

$$\int \cos^4 x\, dx.$$

Note that in order to be able to use an orindary substitution like $u = \sin x$ or $u = \cos x$, we would to have either a single $\cos x$ term or a single $\sin x$ term so that the du term would be correct. In this case using $\sin^2 x + \cos^2 x = 1$ to rewrite the given integrand will not produce such a single $\sin x$ or $\cos x$ term, so instead we must a half-angle formula:

$$\cos^2\theta = \frac{1}{2}(1+\cos 2\theta).$$

Using this once we can rewrite the given integrand as:

$$(\cos^2 x)^2 = \left[\frac{1}{2}(1+\cos 2x)\right]^2 = \frac{1}{4}(1+2\cos 2x+\cos^2 2x),$$

so

$$\int \cos^4 x\,dx = \frac{1}{4}\int (1+2\cos 2x+\cos^2 2x)\,dx.$$

Now, the first two pieces are ones we can integrate (noting that to get integrate $2\cos 2x$ we can use the substitution $u = 2x$), but for the final term we again need to use a half-angle formula. In this case, the role of θ is played by $2x$, so:

$$\cos^2 2x = \frac{1}{2}(1+\cos 4x) = \frac{1}{2}+\frac{1}{2}\cos 4x.$$

Thus

$$\begin{aligned}\frac{1}{4}\int (1+2\cos 2x+\cos^2 2x)\,dx &= \frac{1}{4}\int \left(1+2\cos 2x+\frac{1}{2}+\frac{1}{2}\cos 4x\right)\,dx \\ &= \frac{1}{4}\left(x+\sin 2x+\frac{1}{2}x+\frac{1}{8}\sin 4x\right)+C.\end{aligned}$$

To be clear, to integrate something like $\cos 4x$ we use a substitution $u = 4x$; then $du = 4\,dx$ and $\frac{1}{4}du = dx$, so

$$\int \cos 4x\,dx = \frac{1}{4}\int \cos u\,du = \frac{1}{4}\sin u + C = \frac{1}{4}\sin 2x + C.$$

As a check, we differentiate:

$$\begin{aligned}\frac{1}{4}\left(x+\sin 2x+\frac{1}{2}x+\frac{1}{8}\sin 4x\right)' &= \frac{1}{4}\left(1+2\cos 2x+\frac{1}{2}+\frac{1}{2}\cos 4x\right) \\ &= \frac{1}{4}+\frac{2}{4}\cos 2x+\frac{1}{2}(1+\cos 4x) \\ &= \frac{1}{4}+\frac{2}{4}\cos 2x+\cos^2 2x \\ &= \left(\frac{1}{2}+\frac{1}{2}\cos 2x\right)^2 \\ &= (\cos^2 x)^2 \\ &= \cos^4 x\end{aligned}$$

as expected. Note that at various points we used half-angle formulas to rewrite certain expressions.

Warm-Up 2. We evaluate the integral

$$\int \tan^3 x \sec^3 x\,dx.$$

Here we use the trig identity

$$\sec^2\theta = \tan^2\theta + 1.$$

From this we also get $\sec^2\theta - 1 = \tan^2\theta$, so we rewrite our integrand as:

$$\tan^3 x \sec^3 x = \tan^2 x \tan x \sec^2 x \sec x = (\sec^2 x - 1)\sec^2 x \sec x \tan x,$$

so

$$\int \tan^3 x \sec^3 x \, dx = \int (\sec^2 x - 1)\sec^2 x \sec x \tan x \, dx.$$

The point is that now the substitution $u = \sec x$ works nicely; we get

$$du = \sec x \tan x \, dx,$$

so

$$\begin{aligned}\int (\sec^2 x - 1)\sec^2 x \sec x \tan x \, dx &= \int (u^2 - 1)u^2 \, du \\ &= \int (u^4 - u^2)\, du \\ &= \frac{1}{5}u^5 - \frac{1}{3}u^3 + C \\ &= \frac{1}{5}\sec^5 x - \frac{1}{3}\sec^3 x + C.\end{aligned}$$

As a check we differentiate:

$$\begin{aligned}\left(\frac{1}{5}\sec^5 x - \frac{1}{3}\sec^3 x\right)' &= \sec^4 x(\sec x \tan x) - \sec^2 x(\sec x \tan x) \\ &= \sec^2 x(\sec^2 x - 1)\sec x \tan x \\ &= \sec^2 x \tan^2 x \sec x \tan x \\ &= \sec^3 x \tan^3 x\end{aligned}$$

as expected. Note that we used $\sec^2 x - 1 = \tan^2 x$ in the third line.

Warm-Up 3. Finally we compute

$$\int \cos x \tan^2 x \, dx.$$

Using $\tan^2 x = \sec^2 x - 1$, we have:

$$\int \cos x \tan^2 x \, dx = \int \cos x \sec^2 x - 1)\, dx = \int (\sec x - \cos x)\, dx,$$

where we use the fact that $\sec x = \frac{1}{\cos x}$. Hence

$$\int \cos x \tan^2 x \, dx = \int (\sec x - \cos x)\, dx = \ln|\sec x + \tan x| - \sin x + C.$$

As a check:

$$(\ln|\sec x + \tan x| - \sin x)' = \frac{\sec x \tan x + \sec^2 x}{\sec x + \tan x} - \cos x$$

$$
\begin{aligned}
&= \frac{\sec x(\tan x + \sec x)}{\sec x + \tan x} - \cos x \\
&= \sec x - \cos x \\
&= \frac{\cos x}{\cos^2 x} - \cos x \\
&= \cos x(\sec^2 x - 1) \\
&= \cos x \tan^2 x
\end{aligned}
$$

as expected.

Moral. We've seen many integration techniques so far, and we'll still see some more in the coming days. These techniques might seem at first glance to be wildly different, but there is actually an underlying theme behind all of them: all integration techniques amount to finding ways of rewriting a given integral as a different-looking integral which gives the same result. Indeed, substitution and trig substitution (which we'll cover today) are all about taking a given integral and rewriting it in terms of a new variable instead, and integration by parts, trig integrals, and partial fractions (which we'll talk about next time) are all about writing a given integral as a new integral still in terms of the same variable. Keeping this common theme in mind can help see exactly what these various integration techniques have in common and what the point of them is.

Trigonometric substitution. As mentioned above, the technique of trigonometric substitution is meant to write a given integral in terms of one variable as an integral in terms of a different variable. This technique applies to integrands where we have a portion which reminds us of a certain trig identity; for instance, something like

$$1 - x^2 \text{ reminds us of } 1 - \cos^2 x = \sin^2 x \text{ or } 1 - \sin^2 x = \cos^2 x,$$

something like

$$x^2 - 1 \text{ reminds us of } \sec^2 x - 1 = \tan^2 x,$$

and something like

$$x^2 + 1 \text{ reminds us of } \tan^2 x + 1 = \sec^2 x.$$

Of course, $1 - x^2, x^2 - 1, x^2 + 1$ are not the only types of expressions on which this technique is useful, but they are the most basic cases. The goal is to make a substitution of the form

$$x = \text{some trig function}$$

to rewrite the given integral in another way, thereby by obtaining something simpler to compute. Let's see this in action.

Example 1. We first compute

$$\int \frac{1}{x^2\sqrt{1-x^2}}\,dx.$$

As pointed out above, the key observation is that the $1 - x^2$ term under the square root has a similar-looking form to the $1 - \cos^2 x = \sin^2 x$ trig identity. Because of this, making the substitution $x = \cos\theta$, where θ is some new variable, will help to rewrite the given integral in a simpler form.

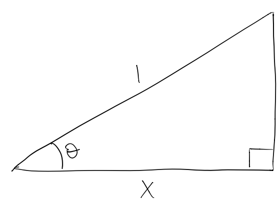
With $x = \cos\theta$, we get $dx = -\sin\theta\,d\theta$. Replacing all instances of x and dx in the original integral gives:

$$\int \frac{1}{x^2\sqrt{1-x^2}}\,dx = \int \frac{1}{\cos^2\theta\sqrt{1-\cos^2\theta}}(-\sin\theta)\,d\theta.$$

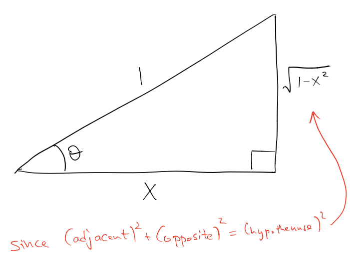
The point is that now we can simplify the square root using $1-\cos^2\theta = \sin^2\theta$, which is the whole reason why we used the substitution we did. We get:

$$\begin{aligned}
-\int \frac{1}{\cos^2\theta\sqrt{1-\cos^2\theta}}\sin\theta\,d\theta &= -\int \frac{\sin\theta}{\cos^2\theta\sqrt{\sin^2\theta}}\,d\theta \\
&= -\int \frac{1}{\cos^2\theta}\,d\theta \\
&= -\int \sec^2\theta\,d\theta. \\
&= -\tan\theta + C.
\end{aligned}$$

Again, note how using the substitution $x = \cos\theta$ allowed us to rewrite the given integral in a form which was possible to compute. But we are not done yet: we must now rewrite the result back in terms of x. In particular, we need to figure out how to express $\tan\theta$ in terms of x. The key is that we know $x = \cos\theta$ due to our original substitution, and this information is enough to figure out what $\tan\theta$ is. Start by drawing a right triangle with a specific angle denoted by θ. Since $x = \cos\theta$ and cosine is "adjacent over hypothenuse", we can label the adjacent sides and hypothenuse by x and 1:



Again, these are the correct values since this indeed gives "adjacent over hypothenuse" equal to $\frac{x}{1} = x$, which is precisely what $\cos\theta$ should be equal to under our substitution. The remaining side can be determined from the Pythagorean Theorem:



The point is that now we can determine $\tan\theta$: since tangent is "opposite over adjacent", we get that

$$\tan\theta = \frac{\sqrt{1-x^2}}{x}.$$

Thus the final answer to our integral is:

$$\int \frac{1}{x^2\sqrt{1-x^2}}\,dx = -\tan\theta + C = -\frac{\sqrt{1-x^2}}{x} + C.$$

As a check, we differentiate our result (using the quotient rule) to see that we get back the original integrand:

$$\left(-\frac{(1-x^2)^{1/2}}{x}\right)' = -\frac{x\frac{1}{2}(1-x^2)^{-1/2}(-2x) - (1-x^2)^{1/2}}{x^2}$$

$$= \frac{\frac{x^2}{\sqrt{1-x^2}} + \sqrt{1-x^2}}{x^2}$$

$$= \frac{\frac{x^2}{\sqrt{1-x^2}} + \frac{1-x^2}{\sqrt{1-x^2}}}{x^2}$$

$$= \frac{1}{x^2\sqrt{1-x^2}},$$

as expected. (Hoorah! Math works!)

Example 2. Finally we compute

$$\int \frac{\sqrt{x^2-4}}{x}\, dx.$$

The form of expression under the square root also suggests a trig substitution should be useful, but in this case we have to be a little careful. If this term was just $x^2 - 1$, then $x = \sec\theta$ is good because this then makes use of the trig identity $\sec^2\theta - 1 = \tan^2\theta$. However, in our case we have $x^2 - 4$, so we need a way to deal with the fact that our constant is 4 instead of 1. But the point is that we can get a constant of 4 by multiplying $\sec^2\theta - 1 = \tan^2\theta$ through by 4:

$$4\sec^2\theta - 4 = 4\tan^2\theta.$$

The left hand side is now precisely what we have under the square root when we use the substitution $x = 2\sec\theta$, since with this substitution we do get $x^2 = 4\sec^2\theta$. The point in this type of substitution is to set

$$x = (\text{constant})(\text{trig function})$$

where the constant is chosen to balance out the constant we have in the original expression.

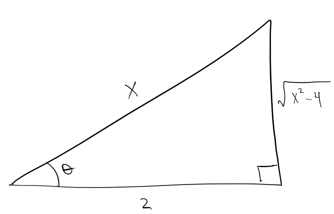
With $x = 2\sec\theta$, we get $dx = 2\sec\theta\tan\theta\, d\theta$, so our integral becomes:

$$\begin{aligned}
\int \frac{\sqrt{x^2-4}}{x}\, dx &= \int \frac{\sqrt{4\sec^2\theta - 4}}{2\sec\theta}(2\sec\theta\tan\theta)\, d\theta \\
&= \int \frac{\sqrt{4\tan^2\theta}}{2\sec\theta}(2\sec\theta\tan\theta)\, d\theta \\
&= \int 4\tan^2\theta\, d\theta.
\end{aligned}$$

We can now compute the resulting integral using the trig identity $\tan^2\theta = \sec^2\theta - 1$:

$$\int 4\tan^2\theta\, d\theta = 4\int (\sec^2\theta - 1)\, d\theta = 4(\tan\theta - \theta) + C.$$

All that remains is to express the result back in terms of x. Since secant is “hypothenuse over adjacent” and $x = 2\sec\theta$ (so $\frac{x}{2} = \sec\theta$), we get the right triangle:



Again, to be clear, $\sec\theta = \frac{x}{2}$ is "hypothenuse over adjacent", which is why the hypothenuse is x and the side adjacent to θ is 2. From this we get that

$$\tan\theta = \frac{\sqrt{x^2-4}}{2}.$$

Finally, since $\frac{x}{2} = \sec\theta$, $\theta = \operatorname{arcsec}(\frac{x}{2})$, so our final answer is:

$$\int \frac{\sqrt{x^2-4}}{x}\,dx = 4\left(\frac{\sqrt{x^2-4}}{2} - \operatorname{arcsec}\frac{x}{2}\right) + C.$$

Note that in this case we cannot express θ is any other way except for $\operatorname{arcsec}\frac{x}{2}$, since the right triangle will only help determine something of the form "trig function evaluated at θ".

Lecture 9: Partial Fractions

Warm-Up 1. We compute the area of the region enclosed by the unit circle, which we know should be π. The unit circle has equation $x^2+y^2=1$, which gives $y = \pm\sqrt{1-x^2}$; the positive square root describes the top half of the unit circle and the negative square root describes the top half. Note that the total area of the region we're loooking at is 4 times the area of the quarter which lies in the first quadrant alone, so the area we want is given by

$$4\int_0^1 \sqrt{1-x^2}\,dx.$$

We use the trig substitution $x = \sin\theta$. Then $dx = \cos\theta\,d\theta$, so

$$4\int_0^1 \sqrt{1-x^2}\,dx = 4\int_0^{\pi/2} \sqrt{1-\sin^2\theta}\cos\theta\,d\theta,$$

where the bounds on the new integral come from finding the values of θ which give the bounds $x=0$ and $x=1$ on the original integral. Computing the resulting integral gives:

$$\begin{aligned}
4\int_0^{\pi/2} \sqrt{\cos^2\theta}\cos\theta\,d\theta &= 4\int_0^{\pi/2} \cos^2\theta\,d\theta \\
&= 4\int_0^{\pi/2} \frac{1}{2}(1+\cos 2\theta)\,d\theta \\
&= 2\left(\theta + \frac{1}{2}\sin 2\theta\right)\Big|_0^{\pi/2} \\
&= 2\left(\frac{\pi}{2}\right) \\
&= \pi
\end{aligned}$$

as expected. Note that in the third step we used the half-angle formula $\cos^2\theta = \frac{1}{2}(1+\cos 2\theta)$.

Warm-Up 2. We now compute the indefinite integral

$$\int \frac{1}{\sqrt{(x-1)^2+16}}\,dx.$$

The key observation is that the integrand involves something of the form

$$(\text{something})^2 + \text{constant}.$$

This should remind us of the trig identity

$$\tan^2\theta + 1 = \sec\theta.$$

We thus use the substitution

$$x - 1 = 4\tan\theta.$$

To be clear, the 4 is there to deal with the fact that constant is $16 = 4^2$ instead of 1 (the point being that squaring $4\tan\theta$ will give the same constant in front as the constant being added afterwards), and the $x - 1$ is there since that is the expression which is being squared in the original integrand. In general, for something of the form

$$(\text{something})^2 + (\text{positive constant}),$$

a substitution where we set

$$\text{something} = \sqrt{\text{positive constant}}\tan\theta$$

would be useful. A similar thing applies to other trig substitutions.

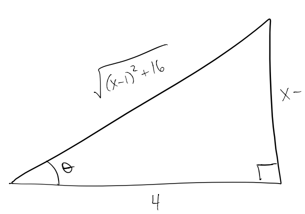
With $x - 1 = 4\tan\theta$, we get

$$dx = 4\sec^2\theta.$$

Thus our integral becomes:

$$\begin{aligned}\int \frac{1}{\sqrt{(x-1)^2+16}}\,dx &= \int \frac{1}{\sqrt{(4\tan\theta)^2+16}}(4\sec^2\theta)\,d\theta \\ &= \int \frac{1}{\sqrt{16(\tan^2\theta+1)}}(4\sec^2\theta)\,d\theta \\ &= \int \frac{1}{4\sec\theta}4\sec^2\theta\,d\theta \\ &= \int \sec\theta\,d\theta \\ &= \ln|\sec\theta + \tan\theta| + C.\end{aligned}$$

Finally, we express the result back in terms of x. Since $x - 1 = 4\tan\theta$, we have $\frac{x-1}{4} = \tan\theta$. Since tangent is "opposite over adjacent", this gives the right triangle:



Thus (since secant is "hypothenuse over adjacent")

$$\sec\theta = \frac{\sqrt{(x-1)^2+16}}{4},$$

so

$$\int \frac{1}{\sqrt{(x-1)^2+16}}\,dx = \ln\left|\frac{\sqrt{(x-1)^2+16}}{4} + \frac{x-1}{4}\right| + C.$$

As usual, as a check you can differentiate this result and verify that you indeed (after a lot of simplifying) get $\frac{1}{\sqrt{(x-1)^2+16}}$ as the derivative.

Partial Fractions. The method of partial fractions allows us to write a fraction of polynomials in a way which makes such an expression simpler to integrate. The type of expression we'll consider is one where we have something like

$$\frac{\text{something}}{(x-a)(x-b)},$$

and the fact that we can factor the denominator as given is what leads to a simpler expression. For instance, we can rewrite $\frac{2}{x^2-1}$ as

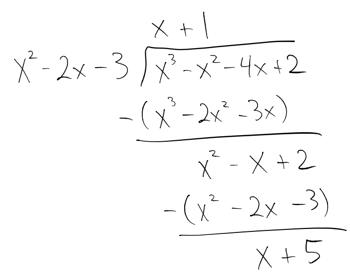
$$\frac{2}{x^2-1} = \frac{2}{(x-1)(x+1)} = \frac{1}{x-1} - \frac{1}{x+1},$$

where the point is that the two terms we end up with are now possible to integrate directly. Let's run through some examples to see how this works.

Example 1. Consider the integral

$$\int \frac{x^3-x^2-4x+2}{x^2-2x-3}\,dx.$$

To simplify the integrand we can first perform polynomial long division; indeed, we must do this first since the method of partial fractions only applies to fractions where the polynomial in the numerator has a (strictly) smaller degree than the one in the denominator, and in this case the numerator has degree 3 and the denominator has degree 2. Long division gives:



(See me in person if you haven't seen long division in a while, or never saw it, and need a refresher!) This says that dividing x^3-x^2-4x+2 by x^2-2x-3 gives $x+1$ with a remainder of $x+5$, so:

$$\frac{x^3-x^2-4x+2}{x^2-2x-3} = x+1+\frac{x+5}{x^2-2x-3}.$$

Thus we get:

$$\int \frac{x^3-x^2-4x+2}{x^2-2x-3}\,dx = \int \left(x+1+\frac{x+5}{x^2-2x-3}\right)dx.$$

Now we work to deal with the

$$\frac{x+5}{x^2-2x-3}$$

term. Since the numerator has smaller degree than the denominator, we can now rewrite this fraction using *partial fractions*. Since the denominator factors as $x^2 - 2x - 3 = (x-3)(x+1)$, the fact is that we can rewrite the given fraction as

$$\frac{x+5}{x^2-2x-3} = \frac{A}{x-3} + \frac{B}{x+1}$$

for an appropriate choice of A and B. Such an expression is called a *partial fraction decomposition.* To find A and B, we first multiplying the given equation through by $x^2 - 2x - 3 = (x-3)(x+1)$ in order to clear denominators. This gives:

$$x + 5 = \frac{A(x-3)(x+1)}{x-3} + \frac{B(x-3)(x+1)}{x+1} = A(x+1) + B(x-3).$$

Note that if we now set $x = 3$ in the resulting equation

$$x + 5 = A(x+1) + B(x-3),$$

we will get an equation which only involves A, and if we set $x = -1$ we get an equation which only involves B; these equations will let us find the precise values of A and B. Setting $x = 3$ gives:

$$3 + 5 = A(3+1) + B(3-3), \text{ or } 8 = 4A, \text{ so } A = 2,$$

and setting $x = -1$ gives:

$$-1 + 5 = A(-1+1) + B(-1-3), \text{ or } 4 = B(-4), \text{ so } B = -1.$$

Thus we get that

$$\frac{x+5}{x^2-2x-3} = \frac{2}{x-3} + \frac{-1}{x+1} = \frac{2}{x-3} - \frac{1}{x+1}$$

is the correct partial fraction decomposition. We can check that this is correct by adding the two fractions on the right using a common denominator:

$$\frac{2}{x-3} - \frac{1}{x+1} = \frac{2(x+1) - 1(x-3)}{(x-3)(x+1)} = \frac{2x+2-x+3}{(x-3)(x+1)} = \frac{x+5}{x^2-2x-3}.$$

Using this partial fraction decomposition we finally arrive at:

$$\int \frac{x^3 - x^2 - 4x + 2}{x^2 - 2x - 3}\, dx = \int \left(x + 1 + \frac{2}{x-3} - \frac{1}{x+1} \right) dx.$$

Now we are left with pieces, each of which we can integrate. We get:

$$\int \frac{x^3 - x^2 - 4x + 2}{x^2 - 2x - 3}\, dx = \frac{1}{2}x^2 + x + 2\ln|x-3| - \ln|x+1| + C$$

as our final answer. (To be clear, we can evaluate $\int \frac{2}{x-3}\, dx$ using the substitution $u = x - 3$, and $\int \frac{1}{x+1}\, dx$ using the substitution $u = x + 1$.) As a check, you can differentiate

$$\frac{1}{2}x^2 + x + 2\ln|x-3| - \ln|x+1|$$

and see that you indeed get $\frac{x^3-x^2-4x+2}{x^2-2x-3}$ as the result after simplifying.

Example 2. We compute the integral

$$\int \frac{e^{4x} - e^{3x} - 4e^{2x} + 2e^x}{e^{2x} - 2e^x - 3}\, dx.$$

As written this might not seem to have anything to do with partial fractions, but the point is that it is if we view it in the right way. In particular, note that each term in the numerator and denominator is $u = e^x$ raised to some power, since $e^{ax} = (e^x)^a$. This suggests that if we use the substitution $u = e^x$, we *will* end up with a partial fractions problem.

With $u = e^x$, we get $du = e^x\, dx$. After we factor a e^x term out of the numerator, we can rewrite the given integral as

$$\int \frac{e^{3x} - e^{2x} - 4e^x + 2}{e^{2x} - 2e^x - 3} e^x\, dx = \int \frac{u^3 - u^2 - 4u + 2}{u^2 - 2u - 3}\, du.$$

Low and behold the integral we get in terms of u is *precisely* the one we computed in Example 1 using partial fractions, only here we are using u instead of x. The result of Example 1 says that:

$$\int \frac{u^3 - u^2 - 4u + 2}{u^2 - 2u - 3}\, du = \frac{1}{2}u^2 + u + 2\ln|u - 3| - \ln|u + 1| + C,$$

and so substituting $u = e^x$ back in gives:

$$\int \frac{e^{4x} - e^{3x} - 4e^{2x} + 2e^x}{e^{2x} - 2e^x - 3}\, dx = \frac{1}{2}e^{2x} + e^x + 2\ln|e^x - 3| - \ln|e^x + 1| + C$$

as the final answer.

This problem illustrate that the method of partial fractions isn't only applicable to integrands which are explicitly given as a fraction of polynomials, but more generally to integrals where we get such a fraction after making an appropriate substitution.

Lecture 10: Approximate Integration

Warm-Up 1. We evaluate the integral

$$\int \frac{e^x}{e^{2x} - 3e^x + 2}\, dx.$$

First, under the substitution $u = e^x$, we get $du = e^x\, dx$ so the given integral becomes

$$\int \frac{e^x}{e^{2x} - 3e^x + 2}\, dx = \int \frac{1}{u^2 - 3u + 2}\, du.$$

Since the denominator of the resulting integrand factors as $u^2 - 3u + 2 = (u - 2)(u - 1)$, we can rewrite the given expression using partial fractions. We setup:

$$\frac{1}{u^2 - 3u + 2} = \frac{A}{u - 2} + \frac{B}{u - 1}.$$

Multiplying through by $(u - 2)(u - 1) = u^2 - 3u + 2$ gives

$$1 = A(u - 1) + B(u - 2),$$

and then setting $u = 1$ gives $B = -1$, and setting $u = 2$ gives $A = 1$. Thus

$$\int \frac{1}{(u-2)(u-1)}\,du = \int \left[\frac{1}{u-2} - \frac{1}{u-1}\right] du = \ln|u-2| - \ln|u-1| + C.$$

Finally, expressing this back in terms of x using $u = e^x$ gives

$$\int \frac{e^x}{e^{2x} - 3e^x + 2}\,dx = \ln|e^x - 2| - \ln|e^x - 1| + C$$

as the final answer.

As a check, we differentiate:

$$\begin{aligned}(\ln|e^x - 2| - \ln|e^x - 1|)' &= \frac{e^x}{e^x - 2} - \frac{e^x}{e^x - 1} \\ &= \frac{e^x(e^x - 1) - e^x(e^x - 2)}{(e^x - 2)(e^x - 1)} \\ &= \frac{e^{2x} - e^x - e^{2x} + 2e^x}{e^{2x} - 3e^x + 2} \\ &= \frac{e^x}{e^{2x} - 3e^x + 2}\end{aligned}$$

as expected.

Warm-Up 2. Now we compute

$$\int \frac{\cos x}{\sin^2 x - 3\sin x + 2}\,dx.$$

Under the substitution $u = \sin x$, so that $du = \cos x\,dx$, the given integral becomes:

$$\int \frac{\cos x}{\sin^2 x - 3\sin x + 2}\,dx = \int \frac{1}{u^2 - 3u + 2}\,du,$$

where the point is that we get the same integral (in terms of u) as in the previous problem. The same partial fractions approach there works, only that now at the end we use $u = \sin x$ to write the result back in terms of x. We get:

$$\int \frac{\cos x}{\sin^2 x - 3\sin x + 2}\,dx = \ln|u-2| + \ln|u-1| + C = \ln|\sin x - 2| + \ln|\sin x - 1| + C$$

as the result.

As a twist, suppose we had an additional factor of $\sin x$ in the numerator of the original integrand:

$$\int \frac{\sin x \cos x}{\sin^2 x - 3\sin x + 2}\,dx.$$

Here the substution $u = \sin x$ gives:

$$\int \frac{u}{u^2 - 3u + 2}\,du.$$

Here we setup the partial fraction decomposition:

$$\frac{u}{u^2 - 3u + 2} = \frac{A}{u-2} + \frac{B}{u-1}.$$

Multiplying through by $u^2 - 3u + 2 = (u-2)(u-1)$ gives

$$u = A(u-1) + B(u-2).$$

Now setting $u = 2$ gives $A = 2$ and setting $u = 1$ gives $B = -1$. Thus

$$\int \frac{u}{u^2 - 3u + 2}\,du = \int \left[\frac{2}{u-2} - \frac{1}{u-1}\right] du,$$

so

$$\int \frac{\sin x \cos x}{\sin^2 x - 3\sin x + 2}\,dx = 2\ln|u-2| - \ln|u-1| + C = 2\ln|\sin x - 2| - \ln|\sin x - 1| + C$$

is our final answer.

Approximating integrals. We have now spent a good amount of time learning how to compute various integrals, and yet we have also pointed out examples of integrals which are not possible to compute directly. For instance,

$$\int_0^1 e^{x^2}\,dx$$

is not possible to compute directly because it is not possible to write down a simple antiderivative of e^{x^2}. (In fact, in a sense which can be made precise, "most" integrals are ones which can't be computed directly, but this is beyond the scope of this course.)

The best we can do with such integrals is approximate their values. We have seen how to do this using rectangles via left/right/midpoint Riemann sum approximations, and now we give two more approximation methods: the *Trapezoid Rule* and *Simpson's Rule*. We'll describe both methods below, and use them to approximate the value of

$$\int_0^1 e^{x^4}\,dx,$$

which is also an integral which cannot be computed directly.

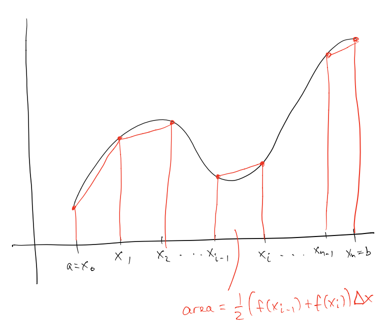
Trapezoid rule. As the name suggests, the Trapezoid Rule uses trapezoids (instead of rectangles) to approximate areas. We take our interval $[a, b]$ and split it up into n subintervals of equal width; we denote the resulting endpoints by

$$a = x_0,\ x_1,\ x_2,\ \ldots,\ x_n = b.$$

The width of each subinterval is denoted Δx (the "change" in x), and is given by

$$\Delta x = \frac{b-a}{n}.$$

The picture behind the Trapezoid Rule is the following:



Over each subinterval, we are looking at a trapezoid whose is that subinterval and whose top edge is given by the line connecting the value of f at the left endpoint to the value of f at the right endpoint. The area of the trapezoid corresponding to the subinterval from x_{i-1} to x_i is

$$\frac{1}{2}(f(x_{i-1}) + f(x_i))\Delta x.$$

We won't dwell on this too much, but you can check the book to see why this is the correct area.

The Trapezoid Rule approximation is then the one obtained by adding up the areas of all of these trapezoids, which gives:

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{2}[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)].$$

The $\frac{1}{2}$ in front comes from the $\frac{1}{2}$ in the formula for the area of each trapezoid. Note the pattern of the expression in brackets: the $f(x_0)$ and $f(x_n)$ terms have a coefficient of 1, but all other intermediate terms have a coefficient of 2. This comes from the fact that, when considering the areas of all trapezoids, the $f(x_0)$ term only appears once as the left endpoint of the first subinterval, the $f(x_n)$ term only appears once as the right endpoint of the final subinterval, but all other intermediate terms appear twice, once as a right endpoint and once as a left endpoint; for instance, x_1 is the right endpoint of the first subinterval but it is also the left endpoint of the second subinterval. For the purposes of this class, knowing this formula without knowing how it is derived is enough.

Example. We use the Trapezoid Rule with $n = 6$ subintervals to approximate the value of

$$\int_0^1 e^{x^4}\,dx.$$

Breaking $[0, 1]$ up into 6 subintervals of equal width gives endpoints:

$$0,\ \frac{1}{6},\ \frac{2}{6},\ \frac{3}{6},\ \frac{4}{6},\ \frac{5}{6},\ 1.$$

The width of each subinterval is

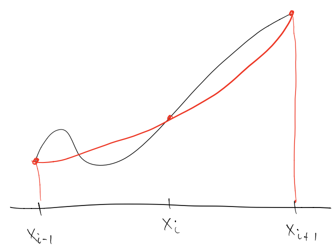
$$\Delta x = \frac{1-0}{6} = \frac{1}{6}.$$

Thus the Trapezoid Rule gives (with $f(x) = e^{x^4}$):

$$\begin{aligned}\int_0^1 e^{x^4}\,dx &\approx \frac{\frac{1}{6}}{2}\left[f(0) + 2f\left(\frac{1}{6}\right) + 2f\left(\frac{2}{6}\right) + 2f\left(\frac{3}{6}\right) + 2f\left(\frac{4}{6}\right) + 2f\left(\frac{5}{6}\right) + f(1)\right] \\ &= \frac{1}{12}\left[e^0 + 2e^{1/6^4} + 2e^{2^4/6^4} + 2e^{3^4/6^4} + 2e^{4^4/6^4} + 2e^{5^4/6^4} + e^1\right].\end{aligned}$$

This is our final answer, but for future reference we note that the expression above is about 1.296.

Simpson's rule. Simpson's Rule uses parabolas to approximate areas. The setup is the same as in the Trapezoid Rule, with one important caveat: Simpson's Rule requires that n, the number of subintervals used, be even. With the same notation Δx and $x_0, x_1, \ldots, x_n$ as before, the picture which lies behind Simpson' Rule is:



The curve in red is the parabola passing through the points on the graph corresponding to x_{i-1}, x_i, x_{i+1}, and the area under this parabola is used in the approximation.

Adding up all of these areas gives:

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{3}[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)].$$

Again, we won't go into why this is the correct sum of areas under parabolas, but you can check the book if interested. Note the pattern of the expression in brackets here: the first coefficient is 1, the final coefficient is 1, and all intermediate coefficients alternate between 4 and 2, starting with 4 and ending with 4 as well. (The fact that n was even guarantees that the there are an odd number of endpoints, which guarantees that the final $f(x_{n-1})$ intermediate term indeed has coefficient 4.)

Back to example. We now use Simpson's Rule with $n = 6$ (even!) subintervals to approximate the value of

$$\int_0^1 e^{x^4}\,dx.$$

We have the same endpoints and Δx as before, so Simpson's Rule gives:

$$\begin{aligned}\int_0^1 e^{x^4}\,dx &\approx \frac{\frac{1}{6}}{3}\left[f(0) + 4f\left(\frac{1}{6}\right) + 2f\left(\frac{2}{6}\right) + 4f\left(\frac{3}{6}\right) + 2f\left(\frac{4}{6}\right) + 4f\left(\frac{5}{6}\right) + f(1)\right] \\ &= \frac{1}{12}\left[e^0 + 4e^{1/6^4} + 2e^{2^4/6^4} + 4e^{3^4/6^4} + 2e^{4^4/6^4} + 4e^{5^4/6^4} + e^1\right].\end{aligned}$$

This is our final answer (we are not expected to be able to simplify something like this), but we note that this value is approximately 1.273.

Recall that the Trapezoid Rule approximation we found before gave an approximate value of 1.296; the *actual* value of this integral (which can only be found indirectly by taking ever better and better approximations) is about 1.271, so in this case Simpson's Rule gives a better approximation. In practical applications, certain times the Trapezoid Rule works better, other times Simpson's Rule, and yet other times rectangles work better. It all depends on the specifics of the particular problem you are looking at.

Fun fact. In the 1990's a medical researcher developed a seemingly new method for estimating the amount of glucose in blood samples, which turned out to be more accurate than existing methods. This was published and lauded as somewhat of a breakthrough at the time. Soon after mathematicians took a look and realized that this researcher had done nothing but rediscover the Trapezoid Rule, which of course had been known to mathematicians for centuries. If nothing else, this speaks to a need for better communication between various scientific communities ;)

Lecture 11: Areas Between Curves, Arclength

Warm-Up. Suppose that the speed at which a car is moving is sampled at different times and results in the following values:

time	1	2	3	4	5
speed	5	10	17	15	12

We estimate the distance that the car traveled between 1 and 5 seconds (which is given by the integral of speed from 1 to 5) using both the Trapezoid Rule and Simpson's Rule with $n = 4$ subintervals. The endpoints we use are thus:

$$x_0 = 1,\ x_1 = 1,\ x_2 = 2,\ x_3 = 3,\ x_4 = 4.\ x_5 = 5,$$

and the length of each subinterval is

$$\Delta x = \frac{5-1}{4} = 1.$$

The Trapezoid Rule gives the approximate distance traveled as:

$$\frac{\Delta x}{2}[f(1) + 2f(2) + 2f(3) + 2f(4) + f(5)] = \frac{1}{2}[5 + 2(10) + 2(17) + 2(15) + 12].$$

Simpson's Rule gives the approximate distance traveled as:

$$\frac{\Delta x}{3}[f(1) + 4f(2) + 2f(3) + 4f(4) + f(5)] = \frac{1}{3}[5 + 4(10) + 2(17) + 4(15) + 12].$$

What do integrals measure? Before moving on we circle back to a fundamental issue, which is understanding what integrals actually do. So far, we have motivated integrals by saying that they measure area, or measure net change when given information about a rate of change. This is all true, but really sells integrals short; if this was all that integrals measured they wouldn't be as useful as they actually are.

It turns out that the uses we've seen thus far are really just instances of a more fundamental idea, that integrals should be interpreted as a type of "infinite sum". To illustrate this we point out one notion from an earlier chapter we skipped over: the average value of a function over an interval. The *average value* of f over $[a, b]$ is given by

$$\frac{1}{b-a}\int_a^b f(x)\,dx.$$

Now, the question is in what sense this gives an average value? When we normally think of an "average" we imagine taking all values we have, add them all together, and then divide by the number of things we have. I claim that this is indeed what the definition above is doing, only accounting for the fact that we have infinitely many values to consider. The integral

$$\int_a^b f(x)\,dx$$

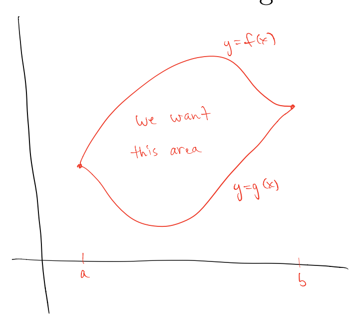
should be interpreted as "adding" together all the values $f(x)$ as x ranges over the interval $[a, b]$. Of course, this is not *literally* true, in particular since it is not so clear what it means to add together infinitely many quantities. But the point is that Riemann sum definition we gave for integrals at the beginning *is* a way to make the idea of infinite summation precise in this setting.

So, going back to the definition of "average value", we are now saying that the integral portion of the definition is indeed the analog of adding together all values $f(x)$ as x ranges throughout $[a,b]$. Now, we claim that the $b-a$ value we are dividing by is the analog of the "number" of values we are adding together. The point is that we can't simply count how many values we have since there are infinitely many, so instead we use the *length* of the interval where the values of x come from as a measure of how "many" such values there are. With this in mind,

$$\frac{1}{b-a}\int_a^b f(x)\,dx$$

is an "infinite" analog of "adding together all values of f and dividing by how many values there are". The moral is that: integration is a type of infinite summation.

Area between curves. We now come to the question of determining the area of a region bounded by two curves. The picture to have in mind is something like:

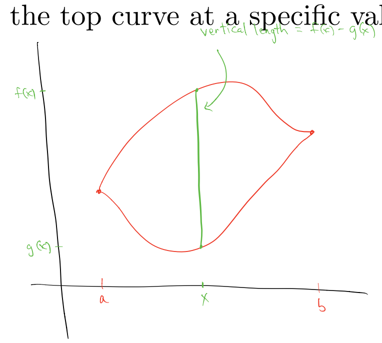


We can obtain the area in question by taking the area under the top curve $y=f(x)$ and subtracting the area under the bottom curve $y=g(x)$. Based on what we know about integrals, the area of the region in question is thus:

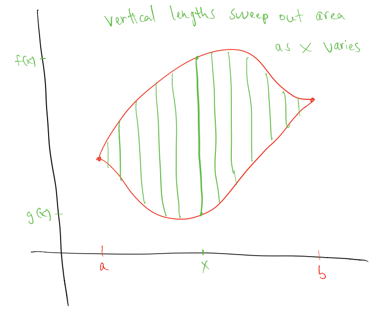
$$(\text{area under } y=f(x)) - (\text{area under } y=g(x)) = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx = \int_a^b [f(x)-g(x)]\,dx.$$

The key point is that we take the integral of the curve on "top" minus the curve on the "bottom".

We'll look at some examples in a second, but first we illustrate how to interpret this formula instead from the "integration is finite summation" point of view mentioned above. From this point of view, we are saying that the area in question is obtained by "adding" together all values $f(x)-g(x)$ as x ranges from $x=a$ to $x=b$. Why does it make sense that this should give the area? We can interpret the difference $f(x)-g(x)$ as the length of the *vertical* line segment which extends from the bottom curve to the top curve at a specific value of x:

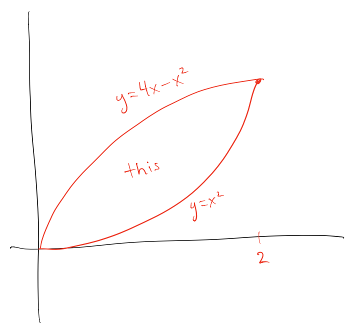


So, the integral $\int_a^b [f(x) - g(x)]\, dx$ is adding together the values of all of these vertical lengths as x ranges through $[a, b]$. The point is that these vertical line segment all together sweep out the region in question:



so that adding together these vertical lengths should indeed give the area of this region! This point of view will be useful in a bit when we look at integrating *with respect to* y instead.

Example 1. We determine the area of the region bounded by the curves $y = x^2$ and $y = 4x - x^2$, which looks something like:



Note that the x-coordinates $x = 0, 2$ specified here are not given in the setup, but are found by finding the points where the two curves intersect. This is done by setting the two y-values equal to each other and solving fo x:

$$x^2 = 4x - x^2, \text{ so } 2x^2 - 4x = 0, \text{ so } 2x(x-2) = 0, \text{ which gives } x = 0 \text{ or } x = 2.$$

These are the values that determine the bounds on the integral we'll need.

From the picture, we can see that $y = 4x - x^2$ is the curve on "top" and $y = x^2$ is the curve on the "bottom". But even if we didn't have a nice picture available, determining which curve is on top and which is below is not so hard: we simply teset the values of the two curves at some intermediate value of x. For instance, $x = 1$ falls without our range $[0, 2]$, and the curve $y = x^2$ gives $y = 1$ at this point while the curve $y = 4x - x^2$ gives $y = 3$. The fact that the second curve gives a larger y value means that it is the curve on "top".

Thus the area of the region in question is:

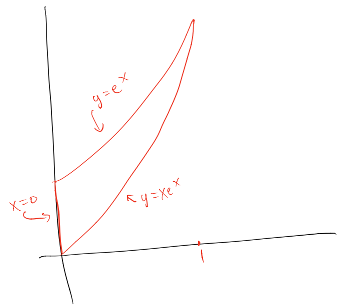
$$\int_0^2 [(4x - x^2) - x^2]\, dx.$$

Computing this gives:

$$\int_0^2 (4x - 2x^2)\, dx = \left(2x^2 - \frac{2}{3}x^3\right)\Big|_0^2 = \frac{8}{3}$$

as the area. Note that, as a sanity check, we should indeed get some positive value since area should be positive; if we had gotten a negative value instead we would know that something went wrong, either in our setup or in the actual integral computation.

Example 2. We determine the area of the region bounded by the curves $y = e^x, y = xe^x$, and $x = 0$, which looks something like:



(Note, however, that technically we don't need the actual picture since we can find intersection points and so on without it. But, a rough picture is indeed a good thing to be able to come up with. This will be more important we when talk about computing *volumes*.) The intersection point in the upper right is found by setting y-values equal to one another:

$$e^x = xe^x, \text{ so } xe^x - e^x = 0, \text{ so } (x-1)e^x = 0, \text{ which gives } x = 1.$$

We thus integrate from $x = 0$ to $x = 1$. (The $x = 0$ value comes from the fact that one of our curves is $x = 0$, which is the y-axis.) Testing $x = \frac{1}{2}$ gives

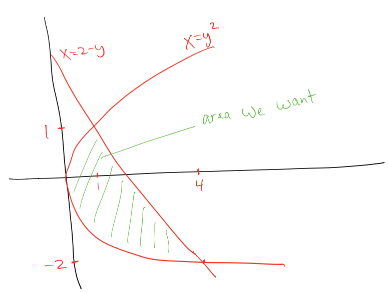
$$e^{1/2} > \frac{1}{2}e^{1/2},$$

so $y = e^x$ is the curve on top. The area of the required region is thus:

$$\int_0^1 [e^x - xe^x]\,dx = \int_0^1 (1-x)e^x\,dx,$$

which can computed using integration by parts with $u = 1 - x$ and $dv = e^x\,dx$. The resulting value is $e - 2$.

Example 3. As a final area example, we determine the area of the region bounded by the curves $x = y^2$ and $x = 2 - y$, which looks like:



The points of intersection come from setting the given x-equations equal to one another:

$$y^2 = 2 - y, \text{ so } y^2 + y - 2 = 0, \text{ so } (y-1)(y+2) = 0, \text{ which gives } y = 1, -2.$$

But now notice that we have a new subtlety: which curve is on "top" depends on where in our given region we're at: between $x = 0$ and $x = 1$, $x = y^2$ is on top, while between $x = 1$ and $x = 4$, $x = 2 - y$ is on top. Because of this, computing this area as an integral with respect to x requires splitting the region into two pieces: the portion before $x = 1$ and the portion after $x = 4$. The area of the former region is

$$\int_0^1 [\sqrt{x} - (-\sqrt{x})]\, dx = \int_0^1 2\sqrt{x}\, dx,$$

where we use $x = y^2$ to get $y = \pm\sqrt{x}$, and note that the positive choice gives the curve on top and the negative choice the curve on the bottom over this particular region. The area of the second region is

$$\int_1^4 [(2 - x) - (-\sqrt{x})]\, dx = \int_1^4 (2 - x + \sqrt{x})\, dx,$$

where we rewrite $x = 2 - y$ as $y = 2 - x$ to get the curve on top and again use that $y = -\sqrt{x}$ is the bottom half of the curve $x = y^2$. Thus the area of the region in question is

$$\int_0^1 2\sqrt{x}\, dx + \int_1^4 (2 - x + \sqrt{x})\, dx.$$

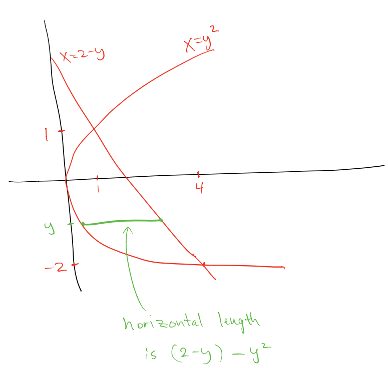
This is certainly possible to now compute, but it was sure a lot of work to get to this point. The issue, again, is that determining which is the curve on "top" depends on where in our region we're at, which requires splitting up the integral. However, note that if instead we consider which is curve on the "left" and which is the curve on the "right", there is not such ambiguity: the curve on the left is $x = y^2$ while the curve on the right is $x = 2 - y$. This suggests that if we instead compute this area using an integral *with respect to* y, it will not be necessary to split up the region into pieces. Indeed, the required area is also given by:

$$\int_{-2}^1 [\text{right} - \text{left}]\, dy = \int_{-2}^1 [(2 - y) - y^2]\, dy.$$

Why should this integral give the correct area and what exactly does it mean to integrate with respect to y?. The answers come from the interpretation of an integral as a sum. The value

$$(2 - y) - y^2$$

gives the length of the *horizontal* line segment extending from the curve on the left to the curve on the right at a specific value of y:



As y ranges over the vertical integral $y = -2$ to $y = 1$, these horizontal segments sweep out the region in question, so adding together their lengths should give the area in question. In this case, since we are adding together values where y ranges from $y = -2$ to $y = 1$, "adding together" should be interpreted as an integral with respect to y, which is what the dy term in the integral indicates. A similar reasoning (that we can compute areas by adding together horizontal lengths as y varies) applies to other regions, and in the end what determines which type of integral (with respect to x or with respect to y) we should use depends on the particulars of the problem at hand; roughly, integrate with respect to x when there is a single "top" curve and a single "bottom" curve, and integrate with respect to y when there is a single "left" curve and a single "right" curve.

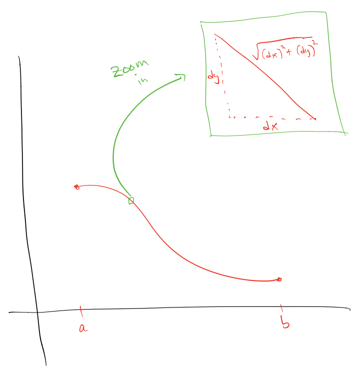
Going back to our example, we now compute:

$$\int_{-2}^{1} [(2-y) - y^2]\,dy = \int_{-2}^{1} [2 - y - y^2]\,dy = \left(2y - \frac{1}{2}y^2 - \frac{1}{3}y^3\right)\Big|_{-2}^{1} = 2 - \frac{1}{2} - \frac{1}{3} - \left(-4 - 2 + \frac{8}{3}\right)$$

as the area of the region in question.

Arclength. The *arclength* of a curve is, as the name suggests, its length. (Imagine "straightening out" the curve so that it becomes a line; the length of this line is the arclength of the curve. Or, imagine walking along the curve; the arclength would be the distance traveled.) For instance, the arclength of a unit circle is its circumference, which is 2π. We are interested in computing the arclength of other curves.

To derive the arclength formula, we again use the interpretation of an integral as a sum. Take a curve $y = f(x)$ and imagining zooming in on an incredibly small portion of it:



In more calculus-like language, we are considering an *infinitesimal portion* of the curve. This infinitesimal portion looks like a line, where we denote the horizontal length by dx (the "infinitesimal" change in x) and the vertical length by dy (the "infinitesimal" change in y). In general, infinitesimal quantities should be thought of as being incredibly small. Note that the dx which shows up in the notation for an integral

$$\int_a^b f(x)\,dx$$

indeed *should* be thought of as an infinitesimal change in x. (The point is that $f(x)\,dx$ should be thought of as the area of an "infinitesimal" rectangle with base length dx and height $f(x)$, and that adding up all of these infinitesimal areas as the integral does should give the area under the graph.)

Going back to the picture above of an infinitesimal portion of our curve, if the horizontal length is dx and the vertical length is dy, the length of this infinitesimal portion (which looks like a line) is given by the Pythagorean Theorem:

$$\sqrt{(dx)^2 + (dy)^2}.$$

Since $y = f(x)$ on our curve, $dy = f'(x)\,dx$ so

$$\sqrt{(dx)^2 + (dy)^2} = \sqrt{(dx)^2 + f'(x)^2(dx)^2} = \sqrt{(1 + f'(x)^2)(dx)^2} = \sqrt{1 + f'(x)^2}\,dx.$$

The conclusion is that the infinitesimal length of a portion of our curve is given by $\sqrt{1 + f'(x)^2}\,dx$, and thus the total length (or arclength) is obtained by adding together all of these infinitesimal lengths as x ranges from $x = a$ to $x = b$, giving

$$\int_a^b \sqrt{1 + f'(x)^2}\,dx$$

as the formula for the arclength of the curve $y = f(x)$ from $x = a$ to $x = b$. Again, this comes about by interpreting an integral as a sum, and by working out what the infinitesimal arclength is.

Infinitesimal quantities. The notion of an infinitesimal quantity helps to make sense of a lot of what we do in calculus. To give one more example of this, note that previously when we used a substitution such as $u = g(x)$ in an integral computation, we had to use

$$du = g'(x)\,dx$$

to figure out how to rewrite dx in terms of du. We never actually explained what was really going on here, but now we can: the equation $du = g'(x)\,dx$ tells us explicitly how to relate an infinitesimal change in u to an infinitesimal change in x, so what we are doing when we make such a substitution is expressing one infinitesimal area

$$f'(g(x))g'(x)\,dx$$

in terms of another infinitesimal area

$$f'(u)\,du.$$

This won't be so important for this course, but this type of manipulation helps to explain many of the formulas involving integrals you might see come up in other courses.

Example. We compute the arclength of the portion of the curve $y = \ln(\cos x)$ which occurs for $0 \le x \le \frac{\pi}{3}$. In this case, $f(x) = \ln(\cos x)$ and

$$f'(x) = \frac{1}{\cos x}\sin x = \tan x.$$

Thus the arclength is given by

$$\int_0^{\pi/3} \sqrt{1 + f'(x)^2}\,dx = \int_0^{\pi/3} \sqrt{1 + \tan^2 x}\,dx.$$

To compute this we use the trig identity $\sec^2 x = 1 + \tan^2 x$ to get:

$$\int_0^{\pi/3} \sqrt{1 + \tan^2 x}\,dx = \int_0^{\pi/3} \sec x\,dx = \ln|\sec x + \tan x|\Big|_0^{\pi/3} = \ln(2 + \sqrt{3}),$$

which is the desired arclength.

Lecture 12: Volumes

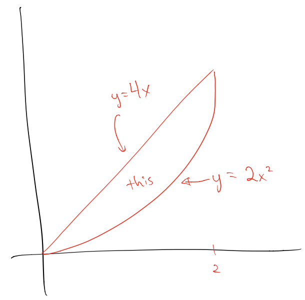
Warm-Up 1. We determine the area of the region bounded by the curves $y = 4x$ and $y = 2x^2$. These curves intersect at values of x where

$$4x = 2x^2, \text{ so } 2x(x-2) = 0, \text{ so at } x = 0, 2.$$

At the intermediate point $x = 1$, the first curves gives $y = 4$ while the second gives $y = 2$, so the first curve is the one on top. The required area is thus:

$$\int_0^2 [4x - 2x^2]\,dx = \left(2x^2 - \frac{2}{3}x^3\right)\Big|_0^2 = \frac{8}{3}.$$

For completeness, the region in question looks like:



Warm-Up 2. We determine the arclength of the portion of the curve $x = y^2 - 2y$ which lies between $y = 0$ and $y = 2$. In this case we setup this arclength as an integral with respect to y. Just to make clear where the formula we'll use comes from, we use the same "infinitesimal length" argument from last time. The length of an infinitesimal portion of this curve is given by

$$\sqrt{(dx)^2 + (dy)^2}.$$

In this case though we have x expressed as a function of y: $x = g(y)$. Thus $dx = g'(y)\,dy$, so

$$\sqrt{(dx)^2 + (dy)^2} = \sqrt{g'(y)^2(dy)^2 + (dy)^2} = \sqrt{g'(y) + 1}\,dy$$

is the infinitesimal length of a piece of the curve. Adding these together as y varies gives:

$$\int_a^b \sqrt{g'(y)^2 + 1}\,dy$$

as the arclength formula.

In our example, $g(y) = y^2 - 2y$, so $g'(y) = 2y - 2$. Hence the arclength of our curve is given by

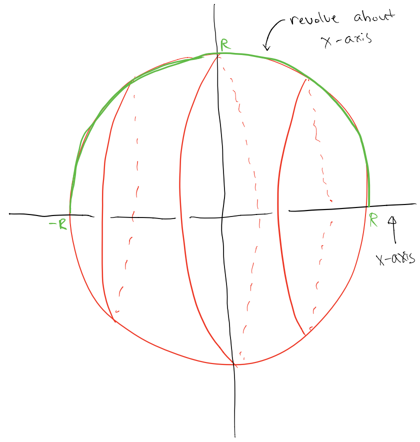
$$\int_0^2 \sqrt{(2y-2)^2 + 1}\,dy.$$

To compute this we can use the trig substitution

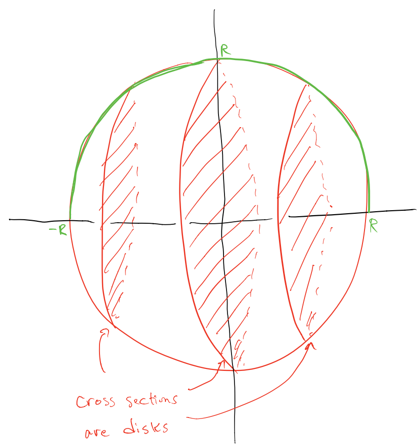
$$2y - 2 = \tan\theta,$$

but I'll leave the details for you to work out. For now we are mainly interested in setting up the appropriate integral.

Volume of a sphere. A well-known formula says that a sphere of radius R has volume $\frac{4\pi R^3}{3}$, but where does this come from? Take the top-half of a circle of radius R and *revolve* it about the x-axis:



The surface obtained by doing this is precisely the sphere of radius R, and this point of view helps us to find the desired volume. If we take the sphere and slice through it with a vertical plane we get something which looks like a *disk*. We refer to this disk as a *cross-section* of the sphere. The point is that if we look at all such cross-sections and add up their areas, we should indeed get the volume of the sphere since these cross-sections sweep out the sphere as x varies:



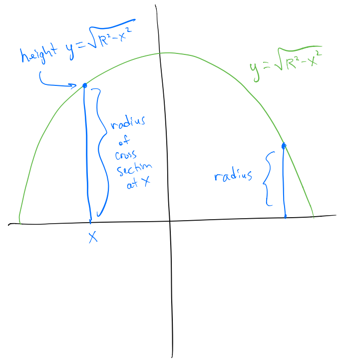
Thus

$$\text{volume of sphere} = \int_{-R}^{R} (\text{area of cross-section})\, dx.$$

Now, since each cross-section is a disk, the area of a cross section is given by

$$\pi(\text{radius})^2.$$

The point is that the radius of the cross-section which occurs at specific value of x is just given by the y-value of the point on the curve $y = \sqrt{R^2 - x^2}$ on the circle which we revolved in order to obtain the sphere:



Thus, the area of this cross-section is

$$\pi(\sqrt{R^2-x^2})^2,$$

so the volume of the sphere is

$$\int_{-R}^{R} \pi(\sqrt{R^2-x^2})^2\,dx = \int_{-R}^{R} \pi(R^2-x^2)\,dx = \pi\left(R^2x - \frac{1}{3}x^3\right)\Big|_{-R}^{R} = \frac{4\pi R^3}{3}$$

as we know it should be.

Volumes in general. The same reasoning applies to other solids. Given a region R in the xy-plane, usually described as being bounded by some curves, we obtain a solid by revolving R around some line. Revolving around a horizontal line where x ranges from $x=a$ to $x=b$ produces a solid whose volume is given by

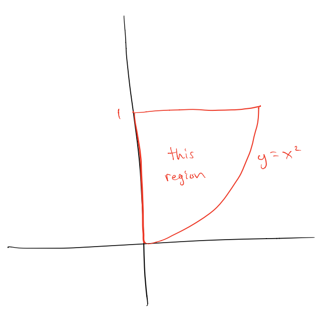
$$\int_a^b (\text{area of cross-section at a specific } x)\,dx$$

while revolving around a vertical line where y ranges from $y=c$ to $y=d$ produces a solid whose volume is given by

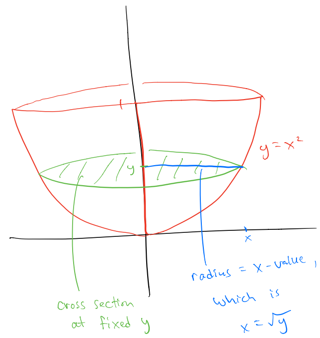
$$\int_c^d (\text{area of cross-section at a specific } y)\,dy.$$

Determining such volumes boil down to figuring out the area of the appropriate cross-sections.

Example 1. Consider the region bounded by the right half of the curve $y=x^2$, the y-axis, and the line $y=1$:



We determine the volume of the solid obtained by revolving this region about the y-axis, which looks something like a bowl which opens upward. A cross-section of this solid looks like a disk of radius $x=\sqrt{y}$, which comes from solving $y=x^2$ for x in terms of y:

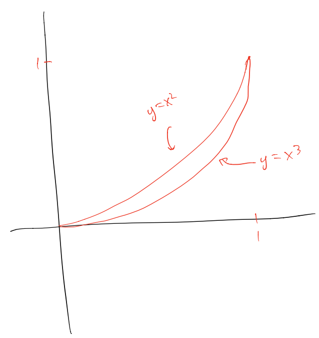


Thus the volume of the solid in question (which is obtained by adding up the areas of the cross-sections as y ranges from $y = 0$ to $y = 1$) is given by:

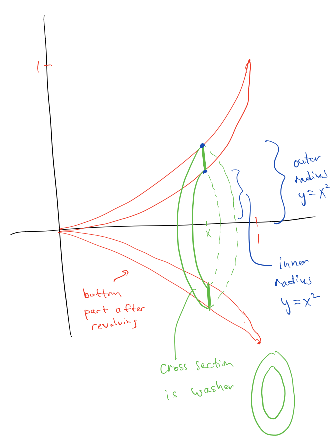
$$\int_0^1 (\text{area of cross-section at a specific } y)\, dy = \int_0^1 \pi(\sqrt{y})^2\, dy = \int_0^1 \pi y\, dy.$$

This integral has value $\frac{\pi}{2}$, so this is the volume of the solid in question.

Example 2. Now consider the region bounded by the curves $y = x^2$ and $y = x^3$, which intersect at $(0, 0)$ and $(1, 1)$. First we determine the volume of the solid obtained by revolving this region around the x-axis. Our picture is:



The difference between this and previous examples is that now cross-sections no longer look like disks, but instead look like *washers*, which are like disks with smaller disks cut out:



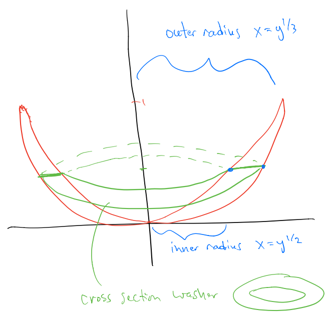
The area of such a cross-section is obtained by taking the area of the larger disk and subtracting the area of the smaller disk:

$$\pi(\text{outer radius})^2 - \pi(\text{inner radius})^2.$$

In our case, the larger disk at a specific value of x has area x^2 (this is the "outer" radius) and the smaller disk has area x^3 (this is the "inner" radius). Thus, the volume of the solid in question is:

$$\begin{aligned}\int_0^1 (\text{area of cross-section at } x)\,dx &= \int_0^1 [\pi(\text{outer radius})^2 - \pi(\text{inner radius})^2]\,dx \\ &= \int_0^1 [\pi(x^2)^2 - \pi(x^3)^2]\,dx \\ &= \pi\int_0^1 [x^4 - x^6]\,dx \\ &= \pi\left(\frac{1}{5} - \frac{1}{6}\right).\end{aligned}$$

Next we take the same region only consider the solid obtained by revolving this region around the y-axis:



Again the (horizontal) cross-sections look like washers, where the outer radius at a specific y is $x = y^{1/3}$ (from $y = x^3$) and the inner radius is $x = y^{1/2}$ (from $y = x^2$). The volume of this solid is thus:

$$\begin{aligned}\int_0^1 (\text{area of cross-section at } y)\,dy &= \int_0^1 [\pi(\text{outer radius})^2 - \pi(\text{inner radius})^2]\,dy \\ &= \int_0^1 [\pi(y^{1/3})^2 - \pi(y^{1/2})^2]\,dy \\ &= \pi\int_0^1 [y^{2/3} - y]\,dy \\ &= \pi\left(\frac{3}{5} - \frac{1}{2}\right).\end{aligned}$$

Again, the key point is that it all comes down to finding the areas of the appropriate cross-sections, which are eitehr washers or disks.

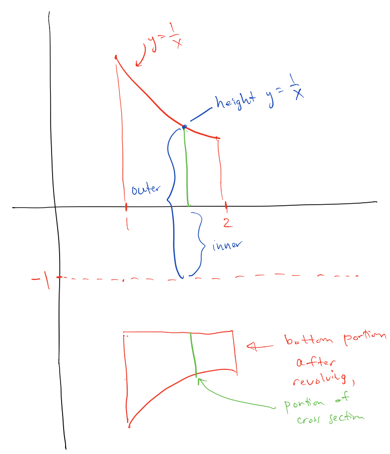
Lecture 13: Improper Integrals

Warm-Up. Consider the region bounded by the curves $y = \frac{1}{x}$ and $y = 0$ between $x = 1$ and $x = 2$:

We determine the volume of the solid obtained by revolving this region about the line $y = -1$. (Note that this is not the x-axis, but instead a different horizontal line.) This volume is given by

$$\int_1^2 (\text{area of cross-section at } x)\, dx.$$

Now, a cross-section here looks like a washer, so we need to determine the inner and outer radii. Here is the key picture:

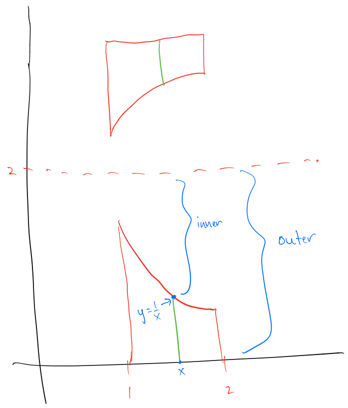


The outer radius takes into account the distance from $y = -1$ up to $y = 0$, and also from $y = 0$ up to $y = \frac{1}{x}$. Thus this outer radius is $1 + \frac{1}{x}$. The innner radius is just the distance from $y = -1$ to $y = 0$, so it is 1. Thus the required volume is given by

$$\int_1^2 [\pi(\text{outer radius})^2 - \pi(\text{inner radius})^2]\, dx = \int_1^2 \pi \left[\left(1 + \frac{1}{x}\right)^2 - 1^2\right] dx.$$

I'll leave it to you to compute this integral because what's new here is setting up the correct integral in the first place.

Imagine that we revolved this same region about the line $y = 2$:



Cross-sections again are washers, but now the outer radius is 2 while the inner radius is $2 - \frac{1}{x}$. Thus the volume of the resulting solid is

$$\int_1^2 \left[\pi(2)^2 - \pi\left(2 - \frac{1}{x}\right)^2\right] dx,$$

which again I'll leave for you to compute. These two examples are trickier than the ones from last time since they require revolving around a line which is not one of the axes, but the method is the same: figure out the area of a cross-section by figuring out the appropriate radius or radii.

Integrals with infinite bounds. As our final integration topic, we consider so-called *improper integrals*, the first type of which are integrals with infinite bounds of integration. To motivate this, consider the the curve $y = \frac{1}{x^2}$ for $x \geq 1$:

The region under the graph of this function extends forever to the right, and yet it turns out that it has a *finite* area. This area is given by the improper integral

$$\int_1^\infty \frac{1}{x^2}\, dx,$$

where the ∞ upper bound indicates an infinitely long interval of integration.

Now, it wouldn't make sense to find the antiderivative of $\frac{1}{x^2}$ and then plug in ∞ and 1 since ∞ is not actual number. (Evaluating a function at ∞ makes no sense.) Instead, we have to be careful as to how such improper integrals are actually defined. We *define* $\int_1^\infty \frac{1}{x^2}\,dx$ to be the value of the limit

$$\lim_{b\to\infty}\int_1^b \frac{1}{x^2}\,dx.$$

This makes sense intuitively: the integral here gives the area of the portion of the region in question which occurs between $x = 1$ and $x = b$, and to get the area of the infinitely long region we want we take this finite area and keep letting b get larger and larger. Thus, to compute such an improper integral, we replace the infinite bound with a new variable, evaluate the result integral, and then take a limit as the new variable goes to ∞.

Example 1. Let us finish the computation of $\int_1^\infty \frac{1}{x^2}\,dx$. Again, technically this integral is defined to be

$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{b\to\infty}\int_1^b \frac{1}{x^2}\,dx.$$

We next compute the integral from 1 to b, all while for now ignoring the limit $\lim_{b\to\infty}$:

$$\lim_{b\to\infty}\int_1^b \frac{1}{x^2}\,dx = \lim_{b\to\infty} -\frac{1}{x}\Big|_1^b = \lim_{b\to\infty}\left(-\frac{1}{b}+1\right).$$

Having computed the integral from 1 to b, we finally take the limit as b goes to ∞:

$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{b\to\infty}\int_1^b \frac{1}{x^2}\,dx = \lim_{b\to\infty}\left(-\frac{1}{b}+1\right) = 1,$$

where we use the fact that $\lim_{b\to\infty}\frac{1}{b} = 0$. Thus

$$\int_1^\infty \frac{1}{x^2}\,dx = 1,$$

so the area under the infinitely long region under the graph of $y = \frac{1}{x}$ for $x \geq 1$ is indeed finite and has value 1.

We say that $\int_1^\infty \frac{1}{x^2}\,dx$ *converges* (or is a convergent integral) since it has a precise, well-defined value. (Or to be more precise, it converges since $\lim_{b\to\infty}\int_1^b \frac{1}{x^2}\,dx$ exists.)

Example 2. Now consider the improper integral

$$\int_1^\infty \frac{1}{x}\,dx$$

This is defined to be

$$\int_1^\infty \frac{1}{x}\,dx = \lim_{b\to\infty}\int_1^b \frac{1}{x}\,dx.$$

Again, we first compute the integral from 1 to b and then take the limit:

$$\lim_{b\to\infty}\int_1^b \frac{1}{x}\,dx = \lim_{b\to\infty} \ln|b|\Big|_1^b = \lim_{b\to\infty}\ln b = \infty.$$

Since $\lim_{b\to\infty}\int_1^b \frac{1}{x}\,dx$ does not exist (saying that the value is infinite does NOT mean it exists), we say that the improper integral $\int_1^\infty \frac{1}{x}\,dx$ *diverges* (or is a divergent integral). Geometrically, this is saying that the area of the infinitely long region under the graph of $y = \frac{1}{x}$ for $x \geq 1$ is infinite.

Important fact which will show up again later. In general, if p denotes some exponent we can consider the integral

$$\int_1^\infty \frac{1}{x^p}\,dx.$$

It turns out that this integral converges only when $p > 1$ (as in Example 1 where $p = 2$), so it diverges when $p \leq 1$ (as in Example 2 where $p = 1$). You should verify this yourself by computing $\int_1^b x^{-p}\,dx$ and then taking the limit as b goes to infinity.

Example 3. We now consider the improper integral

$$\int_{-\infty}^\infty xe^{-x^2}, dx.$$

Now we have two infinite bounds, so the interval of integration is infinitely long in both directions. Again, we have to be careful about how such an integral is actually defined. As a first guess, by using what we did for $\int_1^b$ as an analogy, we might say "replace ∞ with b and $-\infty$ with $-b$ and then take a limit as $b \to \infty$"; that is, we might try to say that the improper integral in question is defined to be

$$\lim_{b\to\infty} \int_{-b}^b xe^{-x^2}\,dx.$$

However, this isn't quite right, or rather, it does happen to work in this particular example but isn't the right thing to do in general. We'll come back to this with an example next time.

To give the *actual* definition of the integral in question, we write it as:

$$\int_{-\infty}^\infty xe^{-x^2}\,dx = \int_{-\infty}^0 xe^{-x^2}\,dx + \int_0^\infty xe^{-x^2}\,dx,$$

and say that it (the expression on the left) exists when *both* improper integrals on the right exist separately. (So, if at least one diverges, the original one would diverge as well.) Each integral on the right is computed by replacing the infinite bound with a new variable (to avoid confusion, I'll use a different new variable for each integral) and then taking a limit:

$$\int_{-\infty}^0 xe^{-x^2}\,dx + \int_0^\infty xe^{-x^2}\,dx = \lim_{b\to-\infty} \int_b^0 xe^{-x^2}\,dx + \lim_{c\to\infty} \int_0^c xe^{-x^2}\,dx.$$

The indefinite integral $\int xe^{-x^2}\,dx$ can be computed using the substitution $u = -x^2$; doing so gives

$$\int xe^{-x^2}\,dx = -\frac{1}{2}e^{-x^2} + C$$

as the result.

Thus, putting it all together, we have:

$$\begin{aligned}
\int_{-\infty}^\infty xe^{-x^2}\,dx &= \int_{-\infty}^0 xe^{-x^2}\,dx + \int_0^\infty xe^{-x^2}\,dx \\
&= \lim_{b\to-\infty} \int_b^0 xe^{-x^2}\,dx + \lim_{c\to\infty} \int_0^c xe^{-x^2}\,dx \\
&= \lim_{b\to-\infty} -\frac{1}{2}e^{-x^2}\Big|_b^0 + \lim_{c\to\infty} -\frac{1}{2}e^{-x^2}\Big|_0^c
\end{aligned}$$

$$= \lim_{b\to-\infty} -\frac{1}{2}(1-e^{-b^2}) + \lim_{c\to\infty} -\frac{1}{2}(e^{-c^2}-1).$$

In the first limit, e^{-b^2} approaches 0 as b approaches $-\infty$, and in the second e^{-c^2} approaches 0 as c approaches ∞, so we get:

$$\int_{-\infty}^{\infty} xe^{-x^2}\,dx = -\frac{1}{2} + \frac{1}{2} = 0$$

as the value of the given improper integral. (So, $\infty_{-\infty}^{\infty} xe^{-x^2}\,dx$ converges.) Note in particular that

$$\int_0^{\infty} xe^{-x^2}\,dx = \frac{1}{2}$$

and

$$\int_0^{\infty} xe^{-x^2}\,dx = -\frac{1}{2},$$

and adding these together is by definition $\int_{-\infty}^{\infty} xe^{-x^2}\,dx$.

Fun fact. The improper integral

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx$$

also converges, and in fact as the value $\sqrt{\pi}$. (The point here is that it is possible to show that this converges without finding the actual value, and you can see this worked out as an example in the book. The precise value of $\sqrt{\pi}$ is not possible to compute using only tools we're learning in this course, but you'll likely see one way of computing it using the notion of a *double integral* in Math 234.)

The fact that

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$$

has important consequences in probability and statistics. Indeed, the graph of the function $f(x) = e^{-x^2}$ models what is usually called a "bell" curve in the theory of so-called normal distributions, and here we are saying that the area of the infinitely long region under this curve is $\sqrt{\pi}$. Good stuff.

Integrals with discontinuities. The second type of improper integral arises when the integrand (the function being integrated) has a discontinuity within the interval of integration. For instance, consider

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx,$$

where the integrand is discontinuous at 0, the left endpoint of the interval $[0, 1]$ of integration. This integral should give the area of the following region, which is infinitely tall:

(The point is that in a definite integral $\int_a^b f(x)\,dx$ of the type we've been considering all along in this course so far, the function f is required to be *bounded* over the entire interval $[a,b]$. This is not the case in the integral we are considering above, which is why this is also an example of an "improper" integral.)

We deal with such an improper integral in the same way we dealt with the previous kind: we replace the troublesome bound with a new variable and then take a limit. In this case, the given improper integral is defined to be

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{b\to 0}\int_b^1 \frac{1}{\sqrt{x}}\,dx.$$

As before, we work out the integral and then take the limit:

$$\lim_{b\to 0}\int_b^1 \frac{1}{\sqrt{x}}\,dx = \lim_{b\to 0} 2x^{1/2}\Big|_b^1 = \lim_{b\to 0} 2(1-\sqrt{b}) = 2.$$

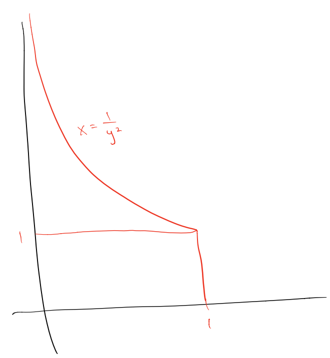
Thus $\int_0^1 \frac{1}{\sqrt{x}}\,dx$ converges and has the value 2.

Note the following subtlety. If we hadn't noticed that the integrand was discontinuous at 0 and had simply tried to compute the given integral using an antiderivative from the beginning we would have gotten the correct answer:

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx \quad \text{"="} \quad 2\sqrt{x}\Big|_0^1 = 2.$$

However, this computation is not valid because of the discontinuity at 0; the fact that this naive attempt just so happens to give the correct value is somewhat of a coincidence and not reflective of a general pattern with such integrals. What makes this integral improper is NOT the issue of plugging in the bounds after we find the antiderivative—it is the discontinuity at zero which makes this not be a definite integral. *Any* integral with a discontinuity is improper and requires rephrasing as a limit.

As another observation, the area in question can also be found using an improper integral with respect to y. Based on the picture



and the fact that $y = \frac{1}{\sqrt{x}}$ means $x = \frac{1}{y^2}$, the given area is also given by

$$1 + \int_1^\infty \frac{1}{y^2}\,dy.$$

The first term of 1 comes from the area of the rectangle at the bottom of the region and then the integral from 1 to ∞ gives the area above this rectangle. Computing this using

$$1 + \lim_{b\to\infty}\int_1^b \frac{1}{y^2}\,dy$$

also gives the value 2 as expected.

Lecture 14: Sequences

Warm-Up 1. We determine whether or not the improper integral

$$\int_{-\infty}^{\infty} x\cos x\,dx$$

converges or diverges. Since we have two infinite bounds, we must approach this integral by considering two separate improper integrals:

$$\int_{-\infty}^{\infty} x\cos x\,dx = \int_{-\infty}^{0} x\cos x\,dx + \int_{0}^{\infty} x\cos x\,dx.$$

We first consider the

$$\int_{-\infty}^{0} x\cos x\,dx = \lim_{b\to-\infty}\int_{b}^{0} x\cos x\,dx$$

portion. The antiderivative of $x\cos x$ can be computed using integration by parts; the result is

$$\int x\cos x\,dx = x\sin x + \cos x + C.$$

Thus

$$\lim_{b\to-\infty}\int_{b}^{0} x\cos x\,dx = \lim_{b\to-\infty}(x\sin x+\cos x)\Big|_{b}^{0} = \lim_{b\to-\infty}(1 - b\sin b - \cos b).$$

This limit does not exist since $b\sin b$ approaches $-\infty$ as b approaches $-\infty$. Thus

$$\int_{-\infty}^{0} x\cos x\,dx$$

diverges, which means that the original integral

$$\int_{-\infty}^{\infty} x\cos x\,dx$$

diverges as well. (Recall that in order for an integral

$$\int_{-\infty}^{\infty} = \int_{-\infty}^{0} + \int_{0}^{\infty}$$

to converge requires that *both* improper integrals on the right converge.

Now, to clarify something we mentioned last time, this example illustrates why trying to define an integral with two infinite bounds as

$$\int_{-\infty}^{\infty} x\cos x\,dx \quad \text{“}=\text{”} \quad \lim_{b\to\infty}\int_{-b}^{b} x\cos x\,dx$$

is a bad idea in general. If you compute

$$\int_{-b}^{b} x\cos x\,dx$$

you'll get a value of 0, so

$$\lim_{b\to\infty}\int_{-b}^{b} x\cos x\,dx = \lim_{b\to\infty} 0 = 0$$

as well. However, this is nonsense: the integral $\int_{-\infty}^{\infty}$ is trying to measure the net area between the graph of $x\cos x$ and the x-axis over the infinitely long interval $(-\infty,\infty)$, and this area should not exist when the half which occurs from $(-\infty,0]$ (as measured by the integral with bounds $\int_{-\infty}^{0}$) does not itself. The fact that the $\lim_{b\to\infty}\int_{-b}^{b}$ integral did give a definite value was somewhat of an accident and doesn't actually represent a definite area. In order for an integral of the form $\int_{-\infty}^{\infty}$ to make sense as a finite value *should* require that both $\int_{-\infty}^{0}$ and $\int_{0}^{\infty}$ exist as finite values independently.

Warm-Up 2. We compute the improper integral

$$\int_1^6 \frac{x}{(2-x^2)^{1/3}}\,dx.$$

The issue here is that the integrand is discontinuous at $x=\sqrt{2}$ where the denominator is zero. Thus, to make sense of this, we must think of this integral as being:

$$\int_1^6 \frac{x}{(2-x^2)^{1/3}}\,dx = \int_1^{\sqrt{2}} \frac{x}{(2-x^2)^{1/3}}\,dx + \int_{\sqrt{2}}^{1} \frac{x}{(2-x^2)^{1/3}}\,dx$$

and then interpret each improper integral on the right as a limit:

$$\int_1^{\sqrt{2}} \frac{x}{(2-x^2)^{1/3}}\,dx = \lim_{b\to\sqrt{2}}\int_1^{b} \frac{x}{(2-x^2)^{1/3}}\,dx \text{ and } \int_{\sqrt{2}}^{6} \frac{x}{(2-x^2)^{1/3}}\,dx = \lim_{cto\sqrt{2}}\int_c^{6} \frac{x}{(2-x^2)^{1/3}}\,dx.$$

The indefinite integral

$$\int \frac{x}{(2-x^2)^{1/3}}\,dx$$

can be computed using the substitution $u=2-x^2$, and you get

$$\int \frac{x}{(2-x^2)^{1/3}}\,dx = -\frac{3}{4}(2-x^2)^{2/3}+C$$

as the result. Thus:

$$\int_1^{\sqrt{2}} \frac{x}{(2-x^2)^{1/3}}\,dx = \lim_{b\to\sqrt{2}}\int_1^{b} \frac{x}{(2-x^2)^{1/3}}\,dx = \lim_{b\to\sqrt{2}} -\frac{3}{4}(2-x^2)^{2/3}\Big|_1^b = \lim_{b\to\sqrt{2}} -\frac{3}{4}[(2-b^2)^{3/2}-1] = 1$$

and

$$\begin{aligned}\int_{\sqrt{2}}^{6} \frac{x}{(2-x^2)^{1/3}}\,dx &= \lim_{c\to\sqrt{2}}\int_c^{6} \frac{x}{(2-x^2)^{1/3}}\,dx \\ &= \lim_{c\to\sqrt{2}} -\frac{3}{4}(2-x^2)^{2/3}\Big|_{\sqrt{2}}^{6} \\ &= \lim_{c\to\sqrt{2}} -\frac{3}{4}[(-34)^{2/3}-(2-c^2)^{2/3}] \\ &= 34^{2/3}.\end{aligned}$$

Hence

$$\int_1^6 \frac{x}{(2-x^2)^{1/3}}\,dx = \int_1^{\sqrt{2}} \frac{x}{(2-x^2)^{1/3}}\,dx + \int_{\sqrt{2}}^{1} \frac{x}{(2-x^2)^{1/3}}\,dx = 1+34^{2/3}.$$

Note that this was somewhat of a lengthy computation (too much work for an exam). I would say that the point here isn't so much getting the actual value, but rather recognizing the steps involved; namely, that this is an improper integral with a discontinuity which requires breaking into two pieces, each of which are meant to be interpreted as limits after we replace the troublesome bound with a new variable.

Warm-Up 3. We show that the integral

$$\int_1^\infty \frac{1}{e^x + x^4}\,dx$$

converges. The new idea here is that we can do this indirectly without carrying about the actual value. Indeed, determining the actual value is not possible since it is not possible to integrate $\frac{1}{e^x+x^4}$ directly. The point is that we can determine converge by *comparing* this integral to another one for which convergence is simpler to determine.

Note that

$$\frac{1}{e^x + x^4} \leq \frac{1}{x^4} \text{ for } x \geq 1.$$

Indeed, the fraction on the right has a smaller denominator than the fraction on the left (since we are adding less), and decreasing a denominator makes a fraction larger. Now, the integral

$$\int_1^\infty \frac{1}{x^4}\,dx$$

is convergent (this is of the form $\int_1^\infty \frac{1}{x^p}\,dx$ with $p > 1$), so I claim this means that the original integral is convergent as well. Indeed, the original integral is definitely positive since the integrand is positive, and the inequality above says that

$$0 \leq \int_1^\infty \frac{1}{e^x + x^4}\,dx \leq \int_1^\infty \frac{1}{x^4}\,dx.$$

Since the integral on the right is finite, the first integral must be as well.

We could have also used the inequality

$$\frac{1}{e^x + x^4} \leq \frac{1}{e^x} \text{ for } x \geq 1,$$

which holds still since the fraction on the right has a smaller denominator than the one on the left, to say that

$$0 \leq \int_1^\infty \frac{1}{e^x + x^4}\,dx \leq \int_1^\infty \frac{1}{e^x}\,dx.$$

The integral on the right converges (meaning is finite, as you can check by computing $\lim_{b\to\infty} \int_1^b e^{-x}\,dx$), so since everyone is positive the first integral must converge as well.

Warm-Up 4. We show that the integral

$$\int_0^1 \frac{e^{2x}(1+\cos x)}{x^2}\,dx$$

diverges. Note that

$$\frac{e^{2x}(1+\cos x)}{x^2} \geq \frac{1}{x^2} \text{ for } 0 < x < 1$$

since the fraction on the right has a smaller numerator. Since everything is positive, this suggests that

$$\int_0^1 \frac{e^{2x}(1+\cos x)}{x^2}\,dx \geq \int_0^1 \frac{1}{x^2}\,dx \geq 0.$$

The improper integral $\int_0^1 \frac{1}{x^2}\,dx$ diverges since

$$\int_0^1 \frac{1}{x^2}\,dx = \lim_{b\to 0}\int_b^1 \frac{1}{x^2}\,dx = \lim_{b\to 0} -\frac{1}{x}\Big|_b^1 = \lim_{b\to 0} -\left(1-\frac{1}{b}\right) = \infty,$$

and thus the larger integral $\int_0^1 \frac{e^{2x}(1+\cos x)}{x^2}\,dx$ must diverge as well.

Summary of comparisons. The key ideas in the last two Warm-Ups are the following: when everything is positive, if the larger one converges, the smaller one must converge as well, and if the smaller one diverges, the larger one must diverge as well. This gives us a way to show an improper integral converges or diverges when there is no hope we'll be able to compute the given integral directly. We'll see a similar idea show up when we discuss convergence of series soon.

Where to next? And so ends the "integration" portion of the course, even though we'll still integrals pop up now and then. We spent a lot of time looking at techniques for computing integrals and a few applications of integrals. The key idea behind it all is that an integral is, essentially, a way to make sense of "adding" together infinitely many quantities; that is, integration is a form of summation.

The remaining portion of the course will be devoted to the notion of a *series*, which is also a type of infinite summation. In the most basic sense, a series is an expression of the form

$$a_1 + a_2 + a_3 + a_4 + \cdots$$

where we add together infinitely many quantities. The $\cdots$ are meant to say "keep going without end". We'll be interested in knowing when such a series converges, which is when such an infinite sum actually gives a finite value as a result. The end goal is to understand the ways in which functions can be approximated using series.

Sequences. Before talking about series, we must first talk about *sequences*, from which series are built. A sequence is nothing but an infinite list of numbers:

$$a_1,\ a_2,\ a_3,\ a_4,\ \ldots.$$

(The connection between sequences and series is that a series is obtained by adding together the terms of a sequence.) To say that a sequence *converges* to a number L is to say that the terms of the sequence get closer and closer to L the further along the sequence we go. Written in terms of limits, a sequence whose n-th term is denoted by a_n converges to a number L when

$$\lim_{n\to\infty} a_n = L.$$

To say that a sequence *diverges* just means that it does not converge.

Example 1. Consider the sequence

$$a_n = 2 + \frac{(-1)^n}{n^2}.$$

To be clear, this notation refers to the sequence whose n-th term is the given a_n. So, the first term of this sequence is teh value a_1 (when $n = 1$), the second term is a_2 (when $n = 2$), and so on. The first few terms of this sequence are thus

$$2 + \frac{-1}{1^2},\ 2 + \frac{1}{2^2},\ 2 + \frac{-1}{3^2},\ 2 + \frac{1}{4^2},\ \ldots.$$

As n goes to infinity, the $\frac{(-1)^n}{n^2}$ gets closer and closer to zero since the numerator bounces back forth between -1 and 1 while the denominator gets larger and larger. Thus

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \left(2 + \frac{(-1)^n}{n^2}\right) = 2 + 0 = 2,$$

so this sequence converges to 2.

Example 2. The sequence

$$b_n = n^2$$

diverges. To be clear, the first few terms of this sequence are

$$1,\ 4,\ 9,\ 16,\ 25,\ \ldots,$$

and the point is that these terms get larger and larger, and thus grow without bound. (We could say they *diverges to infinity.*)

Example 3. The sequence

$$c_n = 1 + (-1)^n$$

also diverges, but for a different reason than that in the previous example. Here the sequence looks like:

$$0,\ 2,\ 0,\ 2,\ 0,\ 2,\ \ldots$$

so this sequence consists of alternating 0's and 2's. Here it's not that these terms get larger and larger, but rather that the alternating behavior prevents them from approaching any one specific value.

Lecture 15: More on Sequences

Warm-Up 1. We determine whether or not the sequence

$$a_n = \frac{3n^2 - 3n}{2n^2 + 4n - 1}$$

converges. As a guess, the key observation is that the n^2 terms in the numerator and denominator "dominate" all the other terms, and so it is these terms that should determine convergence/divergence. Thus, this sequence should behave similarly to the sequence

$$\frac{3n^2}{2n^2} = \frac{3}{2},$$

which converges to $\frac{3}{2}$. In order to make this guess precise we call some things about manipulating limits from Math 220 or an equivalent course. Here we can rewrite the given sequence by dividing the entire numerator and denominator each by n^2. We get:

$$\lim_{n\to\infty} \frac{3n^2 - 3n}{2n^2 + 4n - 1} = \lim_{n\to\infty} \frac{3 - \frac{3}{n}}{2 + \frac{4}{n} - \frac{1}{n^2}}.$$

The $\frac{3}{n}$, $\frac{4}{n}$, and $\frac{1}{n^2}$ terms go to 0 as n goes to infinity, so we get

$$\lim_{n\to\infty} \frac{3-\frac{3}{n}}{2+\frac{4}{n}-\frac{1}{n^2}} = \frac{3-0}{2+0-0} = \frac{3}{2}$$

as the value of the limit. Hence the sequence a_n converges to $\frac{3}{2}$.

A similar technique works for any sequence defined by taking a fraction of polynomials. In general, when the highest power of n in the numerator is greater than that in the denominator, the sequence will diverge; when the highest power of n in the numerator is smaller than that in the denominator, the sequence will converge to 0; and when these highest powers of n are the same, the sequence will converge to the fraction obtained by taking the coefficients of these highest powers. However, you SHOULD work out such limits carefully using the same technique as above where we divide numerator and denominator by a power of n to simplify the given expression.

Warm-Up 2. First we consider the sequence

$$b_n = \frac{1}{2^n}, \text{ which can be written as } b_n = \left(\frac{1}{2}\right)^n.$$

Writing out a few terms, this sequence looks like

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots.$$

The fact that the numerator is staying the same but the denominator is getting larger and larger says that this sequence should converge to 0:

$$\lim_{n\to\infty} \frac{1}{2^n} = 0.$$

Contrast this with the sequence

$$c_n = 2^n,$$

which looks like

$$2,\ 4,\ 8,\ 16,\ 32,\ \ldots.$$

In this case this sequence is increasing without end, so it diverges (to infinity). These examples illustrate a general fact: when r is some number between -1 and 1 (so $-1 < r < 1$), the sequence r^n obtained by taking larger and larger powers of r converges to 0, while when r is a number smaller than -1 or larger than 1 (so $r < -1$ or $1 < r$), the sequence r^n diverges. To finish of the possibilities, $(-1)^n$ diverges and 1^n converges. These are key examples which will play a bigger role when we talk about series.

Example with L'Hopital's rule. Consider the sequence

$$a_n = n^2 e^{-n}, \text{ which can be written as } a_n = \frac{n^2}{e^n}.$$

Note that here both the numerator and denominator are getting larger and larger (so going to ∞) as n goes to infinity. However, the denominator goes to ∞ much faster than the numerator, which suggests the sequence should converge to 0. To make this precise we need to use L'Hopital's rule.

Consider the function

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

The sequence we're looking at is the sequence obtained by plugging in whole numbers n into f. We want to determine

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

Since the numerator and denominator both go to ∞, L'Hopital's rule applies and says that we can attempt to compute the limit by taking the derivative of the numerator and denominator:

$$\lim_{x\to\infty} \frac{x^2}{e^x} = \lim_{x\to\infty} \frac{2x}{e^x}.$$

In this new fraction, the numerator and denominator each again go to ∞, so we can apply L'Hopital's rule to get:

$$\lim_{x\to\infty} \frac{x^2}{e^x} = \lim_{x\to\infty} \frac{2x}{e^x} = \lim_{x\to\infty} \frac{2}{e^x}.$$

This final limit is zero since the numerator has limit 2 and the denominator ∞, so the fraction goes to 0. Hence

$$\lim_{x\to\infty} \frac{x^2}{e^x} = 0.$$

Since the original sequence is obtained by plugging in whole numbers n into $f(x)$, we also get

$$\lim_{n\to\infty} n^2 e^{-n} = 0,$$

so a_n converges to 0.

Squeeze theorem. Consider now the sequence

$$b_n = \frac{n!}{n^n}.$$

Again, both the numerator and denominator are going to infinity, but the fact that the denominator goes to infinity more quickly than the numerator suggests that b_n might converge to zero. However, to justify this in this case we need to recall a fact from Math 220: the *squeeze theorem*.

Note that b_n is always bigger than or equal to zero. If we write out what the numerator and denominator look like, we get:

$$0 \leq \frac{n!}{n^n} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdots n}{n \cdot n \cdot n \cdot n \cdots n}.$$

To be clear, we wrote out $n!$ as "1 times 2 times 3 times $\cdots$ all the way up to n", and we write out n^n has n times itself n times. Note that each term in the numerator is smaller than or equal to the corresponding term in the denominator below it. In particular, in:

$$0 \leq \frac{n!}{n^n} = \frac{1}{n}\left[\frac{2 \cdot 3 \cdot 4 \cdots n}{n \cdot n \cdot n \cdots n}\right],$$

the fraction in brackets is less than 1 since the entire numerator is smaller than the entire denominator. This says that the entire term on the right is less than $\frac{1}{n} \cdot 1 = \frac{1}{n}$, so

$$0 \leq \frac{n!}{n^n} = \frac{1}{n}\left[\frac{2 \cdot 3 \cdot 4 \cdots n}{n \cdot n \cdot n \cdots n}\right] \leq \frac{1}{n}.$$

The terms 0 on the left converge to 0 and the term $\frac{1}{n}$ on the right converges to 0, so the *squeeze theorem* says that the term $b_n = \frac{n!}{n^n}$ in the middle converges to 0 as well.

In general, the squeeze theorem says that when $a_n \leq b_n \leq c_n$ and the sequences a_n and c_n each converge to the *same* value, the sequence b_n converges to that value as well.

Another example. Here is another squeeze theorem example. Consider the sequence

$$a_n = \frac{2 + \sin n}{n}.$$

Since $\sin n$ is always between -1 and 1, we have the inequalities

$$\frac{2 + (-1)}{n} \leq \frac{2 + \sin n}{n} \leq \frac{2 + 1}{n}.$$

The sequence on the left is $\frac{1}{n}$ and converges to 0, and the sequence on the right is $\frac{3}{n}$, which also converges to 0. Hence the squeeze theorem says that $a_n = \frac{2+\sin n}{n}$ converges to 0 as well.

From this we can also look at the sequence

$$b_n = \ln(2 + \sin n) - \ln n, \text{ which can be written as } b_n = \ln\left(\frac{2 + \sin n}{n}\right).$$

We just saw above that $\frac{2+\sin n}{n}$ converges to 0. As the input into the $\ln x$ function goes to 0, $\ln x$ itself goes to $-\infty$, so our sequence b_n diverges.

Increasing, decreasing, and bounded sequences. Finally, consider the sequence

$$a_n = \frac{2n - 3}{3n + 4}.$$

Using previous techniques we can show that this converges to $\frac{2}{3}$, but that is not the point here. Writing out the first few terms of this sequence gives:

$$-\frac{1}{7}, \frac{1}{10}, \frac{3}{13}, \frac{5}{16}, \ldots,$$

and the observation is that it appears as if this sequence is *increasing*, meaning that each term is larger than the one which came before it. To verify that this is indeed the case (since we very well can't actually write down *all* terms in this sequence and verify that each is larger than the one before it), we consider the function

$$f(x) = \frac{2x - 3}{3x + 4}.$$

To show that this function is increasing we can check that its derivative is positive! We have:

$$f'(x) = \frac{(3x + 4)2 - (2x - 3)3}{(3x + 4)^2} = \frac{17}{(3x + 4)^2},$$

which is indeed always positive. Hence f is increasing, so a_n (which is obtained by plugging whole numbers n into f) is increasing as well.

A second observation is that this sequence is *bounded*, which means that all its values lie within some bounded interval. Since this sequence is increasing we know that the first term $-\frac{1}{7}$ is smaller than all other terms, and we also have:

$$\frac{2n - 3}{3n + 4} \leq \frac{2n}{3n} = \frac{2}{3}$$

by making the numerator larger and the denominator smaller. Thus all terms in this sequence satisfy

$$-\frac{1}{7} \leq a_n \leq \frac{2}{3},$$

which shows that a_n is bounded. The point is that *any* sequence which is increasing *and* bounded will always converge, and similarly any sequence which is decreasing (meaning each term is smaller than the one before it) and bounded will converge. This gives a way to show that certain sequences converge without having to worry about computing the limit. This fact won't be so important in our course, but is actually crucial in more theoretical areas.

Lecture 16: Series

Warm-Up 1. We verify that the following sequence converges:

$$a_n = \frac{2 \cdot 3^{n+1}}{5^n} + \frac{\sin(n^2) + n!}{(3n)^n + 1}.$$

We consider each summand separately. The first piece can be written as:

$$\frac{2 \cdot 3^{n+1}}{5^n} = \frac{2 \cdot 3^n \cdot 3}{5^n} = 6\left(\frac{3}{5}\right)^n.$$

The point of writing it like this is that now we have a portion of the form

$$(\text{number})^n,$$

and since this "number" $\frac{3}{5}$ in this case is less than 1, this portion converges to 0. Hence

$$\lim_{n\to\infty} \frac{2 \cdot 3^{n+1}}{5^n} = \lim_{n\to\infty} 6\left(\frac{3}{5}\right)^n = 6 \cdot 0 = 0.$$

Now, break up the second summand into two pieces:

$$\frac{\sin(n^2) + n!}{(3n)^n + 1} = \frac{\sin(n^2)}{(3n)^n + 1} + \frac{n!}{(3n)^n + 1}.$$

For the first, we use the fact that $\sin(n^2)$ is between -1 and 1 to say:

$$-\frac{1}{(3n)^n + 1} \leq \frac{\sin(n^2)}{(3n)^n + 1} \leq \frac{1}{(3n)^n + 1}.$$

The term on the left and the term on the right both converges to 0 since the denominators go to ∞, so the term in the middle also converges to 0.

Finally, we have

$$0 \leq \frac{n!}{(3n)^n + 1} \leq \frac{n!}{(3n)^n} = \frac{n!}{3^n n^n} = \left(\frac{1}{3}\right)^n \frac{n!}{n^n},$$

where the second inequality came from making the denominator smaller. This final term on the right converges to 0 (recall we showed last time that $\frac{n!}{n^n}$ converges to 0), so the squeeze theorem says that

$$\frac{n!}{(3n)^n + 1}$$

converges to zero as well. To sum up, we have shown that each piece of

$$a_n = \frac{2 \cdot 3^{n+1}}{5^n} + \frac{\sin(n^2)}{(3n)^n + 1} + \frac{n!}{(3n)^n + 1}$$

converges to 0, so a_n converges to 0.

Warm-Up 2. This Warm-Up is more conceptual, and is meant to highlight a subtle point. The problem is to come up with two sequences a_n and b_n, neither of which converge, but for which $a_n + b_n$ does converge. One possible example is

$$a_n = (-1)^n \text{ and } b_n = -(-1)^n.$$

Neither of these converges but their sum is $a_n + b_n = 0$, which does converge. Another example is

$$a_n = \cos n + \frac{1}{n} \text{ and } b_n = -\cos n.$$

Again neither of these converge (the $\cos n$ term is what causes both to diverge), but their sum is $a_n + b_n = \frac{1}{n}$, which does converge. There are plenty of more examples you can come up with.

The upshot is that we have to be careful when trying to add sequences: it is true that if a_n and b_n both converge, $a_n + b_n$ will converge as well, but just because $a_n + b_n$ converges does not guarantee that a_n and b_n each converge.

Series. Recall that a *series* is an infinite sum, i.e. an expression where we attempt to add together infinitely many quantities. (To be clear, what makes this as "infinite" sum is the fact that we are adding infinitely many things, NOT that the resulting sum itself might be infinite.) We use the same $\sum$ notation for series we used for Riemann sums, only now indicate the fact that our sums go on forever without end. To be clear, the notation

$$\sum_{n=0}^{\infty} a_n$$

denotes the infinite sum obtained by adding together all terms of the sequence a_n, starting at $n = 0$ and going beyond. So, in this case we get

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + a_3 + a_4 + \cdots,$$

where the $\cdots$ indicate that our sum is without end. Series don't have to start at 0; for instance,

$$\sum_{n=2}^{\infty} a_n = a_2 + a_3 + a_4 + \cdots$$

begins the infinite sum at $n = 2$ instead.

The key question we care about is whether a series *converges*, meaning that we actually get a finite value out of the given infinite sum, or *diverges*, meaning that we don't get a specific value. It is kind of amazing that even though we are adding together infinitely many quantities, we often get finite sums as a result.

Example 1. Consider the series

$$\sum_{n=1}^{\infty} n^2.$$

The first term when $n = 1$ is $1^2 = 1$, the second term when $n = 2$ is $2^2 = 4$, and so on. Writing out this series as an infinite sum gives

$$\sum_{n=1}^{\infty} n^2 = 1 + 4 + 9 + 16 + \cdots .$$

In this case, this infinite sum does not result in a finite value. One way to say this is that the terms we are adding on at each step are getting larger and larger, which in turn makes the resulting sum larger and larger. Another way to say this is to note that each term in our sum bigger than or equal to 1, so this given sum should be larger than the sum obtained by replacing each term by 1:

$$1 + 4 + 9 + 16 + \cdots \geq 1 + 1 + 1 + 1 + \cdots .$$

But adding together infinitely many 1's certainly results in ∞ as the value, so our sum, which is larger, should be infinite as well and hence does not converge, so it diverges.

First divergence test. The previous example already illustrates our first divergence test: in a series $\sum_{n=1}^{\infty} a_n$, if the *sequence* does not converge to 0, the *series* $\sum_{n=1}^{\infty} a_n$ must diverge. Now, it is absolute crucial here to recognize what this is saying, which highlights the difference between sequences and series. Saying that the *series* $\sum_{n=1}^{\infty} a_n$ converges is NOT the same as saying that the sequence a_n converges! These are related concepts for sure, but they do not mean the same thing. (This is a very common cause of confusion.) The *sequence* a_n describes the individual terms being added together to produce the *series* $\sum a_n$.

What this first divergence test says is that if the terms of the sequence a_n do not themselves approach 0, the infinite sum

$$a_1 + a_2 + a_3 + a_4 + \cdots$$

cannot possibly exist. In the previous example, the sequence n^2 does not converge to 0 (in fact it diverges), so the series $\sum_{n=1}^{\infty} n^2$ cannot converge. The intuitive idea is that in order for an infinite sum

$$a_1 + a_2 + a_3 + a_4 + \cdots$$

to have any hope of resulting in a finite value, it had better be the case that the terms being added on at each step are getting smaller and smaller; if this is not the case, the series must diverge.

WARNING. So, possibly the first thing to do when considering whether a series $\sum a_n$ converges or diverges is to see what is happening with the sequence a_n: if $\lim_{n\to\infty} a_n \neq 0$, you're finished—the series $\sum a_n$ will diverge. But as another warning (also a common cause of confusion): just because $\lim_{n\to\infty} a_n$ does equal 0 does NOT mean that the corresponding series $\sum a_n$ converges! Knowing that a_n converges to 0 only tells us that the series $\sum a_n$ has some hope of converging, but does not by itself tell us that the series does indeed converge; this is why we need to consider further convergence tests, as we'll develop in the coming days. Again, the convergence of the *sequence* a_n is related to the convergence of the series $\sum a_n$, but it is not literally the same idea; sequences and series are related but different concepts!

Geometric series. One of the most important types of series is what is known as a *geometric* series. This is a series of the form

$$\sum_{n=0}^{\infty} r^n,$$

where we are adding together higher and higher powers of a number r. So, written as an infinite sum, a geometric series looks like

$$\sum_{n=0}^{\infty} r^n = 1 + r + r^2 + r^3 + r^4 + \cdots .$$

The first term of 1 comes from $r^0 = 1$. Geometric series are important because it is often the case that other series can be related to or compared with such a series, which is good because we can tell exactly when a geometric series *and* what it converges to when it does!

Here is the basic fact you should know by heart: when r is a number outside of the interval $(-1, 1)$ (so when $r \leq -1$ or $1 \leq r$), the geometric series $\sum_{n=0}^{\infty} r^n$ diverges, while when r is a number in the interval $(-1, 1)$ (so when $-1 < r < 1$), the geometric series $\sum_{n=0}^{\infty} r^n$ converges to the value $\frac{1}{1-r}$. We write this as

$$\sum_{n=0}^{\infty} r^n = 1 + r + r^2 + r^3 + \cdots = \frac{1}{1-r} \text{ when } -1 < r < 1.$$

Amazingly this is saying that if we add up the numbers $1, r, r^2, r^3, r^4, \ldots$ we get $\frac{1}{1-r}$ as the result, at least when $-1 < r < 1$, even though there are infinitely many numbers we are adding together. So, this is a concrete example of an infinite sum which actually has a finite value.

Example 2. We show that the series

$$\sum_{n=0}^{\infty} \frac{2 \cdot 3^{n+1}}{5^n}$$

converges and determine its value. The key is that we can express this series in terms of a geometric series. By writing the terms we are adding together, we get:

$$\sum_{n=0}^{\infty} \frac{2 \cdot 3^{n+1}}{5^n} = \sum_{n=0}^{\infty} \frac{2 \cdot 3^n \cdot 3}{5^n} = \sum_{n=0}^{\infty} 6\left(\frac{3}{5}\right)^n .$$

After "factoring out" the 6, we're left with a geometric series with $\frac{3}{5} < 1$, so:

$$\sum_{n=0}^{\infty} 6\left(\frac{3}{5}\right)^n = 6\sum_{n=0}^{\infty} \left(\frac{3}{5}\right)^n = 6\frac{1}{1-\frac{3}{5}},$$

where we use

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

when $r = \frac{3}{5}$ to get the value $\frac{1}{1-\frac{3}{5}}$ of $\sum_{n=0}^{\infty} \left(\frac{3}{5}\right)^n$. Thus the original series converges to 15:

$$\sum_{n=0}^{\infty} \frac{2 \cdot 3^{n+1}}{5^n} = 6\sum_{n=0}^{\infty} \left(\frac{3}{5}\right)^n = 6\frac{1}{1-\frac{3}{5}} = 6\frac{1}{\frac{2}{5}} = \frac{30}{2} = 15.$$

To clarify one point: why is that we can "factor out" 6 as we did above? The point is that this is simply a version of the *distributive* property of multiplication. If we write out the terms of the sum

$$\sum_{n=0}^{\infty} 6\left(\frac{3}{5}\right)^n$$

we get:

$$6+6\left(\frac{3}{5}\right)+6\left(\frac{3}{5}\right)^2+6\left(\frac{3}{5}\right)^2+\cdots.$$

Here we can factor the 6 out to get:

$$6\left(1+\frac{3}{5}+\frac{3^2}{5^2}+\frac{3^3}{5^3}+\cdots\right),$$

which can be written back in terms of σ notation as:

$$6\sum_{n=0}^{\infty}\left(\frac{3}{5}\right)^n.$$

The point, again, is that this is just an infinite version of the distributive property.

Alternatively, imagine we had the same series only with a different starting point:

$$\sum_{n=1}^{\infty}6\left(\frac{3}{5}\right)^n.$$

How do we determine the value of this series? If we write out these terms we get:

$$6\left(\frac{3}{5}\right)+6\left(\frac{3}{5}\right)^2+6\left(\frac{3}{5}\right)^3+\cdots.$$

The key observation is that this sum is the same as the one we had above:

$$6+6\left(\frac{3}{5}\right)+6\left(\frac{3}{5}\right)^2+6\left(\frac{3}{5}\right)^2+\cdots$$

only that the initial 6 (the zeroth) term is missing. So, this new sum should be equal to the old one minus 1:

$$\sum_{n=1}^{\infty}6\left(\frac{3}{5}\right)^n=\sum_{n=0}^{\infty}6\left(\frac{3}{5}\right)^n-1.$$

We found the value of the old one above to be 15, so the value of this new series is 24:

$$\sum_{n=1}^{\infty}6\left(\frac{3}{5}\right)^n=\sum_{n=0}^{\infty}6\left(\frac{3}{5}\right)^n-1=25-1=24.$$

In general, since

$$\sum_{n=0}^{\infty}a_n=a_0+a_1+a_2+\cdots=a_0+\sum_{n=1}^{\infty}a_n,$$

we get

$$\sum_{n=1}^{\infty}=\sum_{n=0}^{\infty}a_n-a_0.$$

A similar thing works for other sums that start at a value larger than 1; for instance:

$$\sum_{n=2}^{\infty}=\sum_{n=0}^{\infty}a_n-a_0-a_1.$$

What does it mean for a series to converge? The question remains: how do we know that $\frac{1}{1-r}$ is the correct value of the series $\sum_{n=0}^{\infty} r^n$? The answer requires us to be a little more precise about what it means for a series to converge. Intuitively, for this particular series to converge means that adding together the infinitely many quantities

$$1 + r + r^2 + r^3 + r^4 + \cdots$$

should give a finite value, but how do we make sense of this when we can't literally sit down and add up all of these infinitely many numbers?

Here is the correct approach. Consider not the infinite sum for now, but rather the sum of the first how ever many terms:

$$1 + r + r^2 + r^3 + \cdots + r^k.$$

This expression is called the *k-th partial sum* of the series $\sum_{n=0}^{\infty} r^n$. So, the zeroeth partial sum is

$$1,$$

the first partial sum is

$$1 + r,$$

the second partial sum is

$$1 + r + r^2,$$

and so on: to get the next partial sum we add on one more term from our series. Intuitively, it should be the case that if the infinite sum

$$1 + r + r^2 + r^3 + \cdots$$

actually exists, the partial sums we're looking at should be getting closer and closer to the value of this infinite sum, since at each step we add on one more term we need. (In a sense, the series itself would be like the "∞-th partial sum", although this is not standard terminology.) Saying the partial sums are getting closer and closer to infinite sum is saying that the *sequence* obtained by taking partial sums converges, and that what they converge to *is* the value of the infinite sum. This leads us to our actual definition of series convergence: we say that $\sum_{n=0}^{\infty} a_n$ converges to L if the sequence of partial sums, whose k-th term is

$$a_0 + a_1 + a_2 + \cdots + a_k,$$

itself converges to L. Again, this is just the precise way of expressing the intuitive thought that in order for an infinite sum to equal some finite value, the finite sums obtained by adding on one more term at each step should themselves be approaching that value. (Think of this as analogous to how Riemann sums approximate integrals, and the actual values of integrals are obtained by taking a limit of Riemann sums.)

For most purposes, this is not a definition we actually have to work with, since we'll have various convergence tests we can use to determine convergence without having resort to the partial sum definition. However, this is the definition which allows us to explicitly determine values of series. Coming back to our example, the k-th partial sum of $\sum_{n=0}^{\infty} r^n$ is

$$1 + r + r^2 + \cdots + r^k.$$

There is a simpler way of expressing such a sum. For instance, when $k = 1$ we have

$$1 + r = \frac{1 - r^2}{1 - r},$$

which we can see by factoring $1-r^2$ into $(1-r)(1+r)$. When $k=2$ we have

$$1+r+r^2=\frac{1-r^3}{1-r},$$

which can see by noting that $(1+r+r^2)(1-r)=1-r^3$. In general, the k-th partial sum can be written as

$$1+r+r^2+\cdots+r^k=\frac{1-r^{k+1}}{1-r}.$$

Now we're in business! Recall that r was a number between -1 and 1. For such numbers, we have

$$\lim_{k\to\infty} r^{k+1}=0.$$

Thus the limit of the k-th partial sums as k goes to ∞ is:

$$\lim_{k\to\infty}\frac{1-r^{k+1}}{1-r}=\frac{1-0}{1-r}=\frac{1}{1-r}.$$

But above we *defined* the value of a series to be the limit of the partial sums, so we get that

$$\sum_{n=0}^{\infty} r^n=\lim_{k\to\infty}(k\text{-th partial sum})=\frac{1}{1-r}$$

as desired. This example was simply to illustrate where $\frac{1}{1-r}$ comes from and is not something you would have to be able to reproduce yourself, but recognizing that a series converges when its sequence of partial sums converges (by definition of what it means for a series to converge) is an important takeaway.

Lecture 17: Integral Test

Warm-Up 1. Consider the series

$$\sum_{n=1}^{\infty}(-1)^n\frac{4^{n+2}}{3^{n-1}}.$$

We claim that this diverges, which we can show by expressing it in terms of a geometric series. We can rewrite the n-term in the sequence being summed up as:

$$(-1)^n\frac{4^{n+2}}{3^{n-1}}=\frac{(-1)^n4^n4^2}{3^n3^{-1}}=48\left(-\frac{4}{3}\right)^n.$$

To be clear, the 48 came from $\frac{4^2}{3^{-1}}=16\cdot 3$. Thus our given series is

$$\sum_{n=1}^{\infty}(-1)^n\frac{4^{n+2}}{3^{n-1}}=\sum_{n=1}^{\infty}48\left(-\frac{4}{3}\right)^n=48\sum_{n=1}^{\infty}\left(-\frac{4}{3}\right)^n.$$

But now the resulting geometric series $\sum_{n=1}^{\infty}\left(-\frac{4}{3}\right)^n$ diverges since $-\frac{4}{3}$ falls outside the interval $(-1,1)$, so the original series diverges as well.

Warm-Up 2. Now consider the series

$$\sum_{n=2}^{\infty}(-1)^n\frac{3\cdot 4^{n-1}}{5^{n+3}}.$$

This converges, which we can determine by writing it in terms of a geometric series; this will also give us a way to find the actual value of this infinite sum. We have:

$$\sum_{n=2}^{\infty}(-1)^n\frac{3\cdot 4^{n-1}}{5^{n+3}} = \sum_{n=2}^{\infty}(-1)^n\frac{3\cdot 4^n\cdot 4^{-1}}{5^n5^3} = \frac{3}{4\cdot 5^3}\sum_{n=2}^{\infty}\left(-\frac{4}{5}\right)^n.$$

Since $-1 < -\frac{4}{5} < 1$, this resulting geometric series converges, so our original one does as well.

Now, we know that

$$\sum_{n=0}^{\infty}\left(-\frac{4}{5}\right)^n = \frac{1}{1-(-\frac{4}{5})} = \frac{5}{9}.$$

Since

$$\sum_{n=0}^{\infty}\left(-\frac{4}{5}\right)^n = 1-\frac{4}{5}+\sum_{n=2}^{\infty}\left(-\frac{4}{5}\right)^n,$$

we get

$$\sum_{n=2}^{\infty}\left(-\frac{4}{5}\right)^n = \sum_{n=0}^{\infty}\left(-\frac{4}{5}\right)^n - 1 + \frac{4}{5}.$$

Thus

$$\frac{3}{4\cdot 5^3}\sum_{n=2}^{\infty}\left(-\frac{4}{5}\right)^n = \frac{3}{600}\left(\frac{5}{9}-1+\frac{4}{5}\right)$$

is the value of our given series.

Warm-Up 3. Finally, we determine whether or not

$$\sum_{n=1}^{\infty}\left(\frac{1}{n}-\frac{1}{n+1}\right)$$

converges. Recall that the precise definition of series convergence involves the notion of a *partial sum*, and the point is that here we can determine the partial sums explicitly. The k-th partial sum is the sum obtained by adding up the terms of the series only up to the $n = k$ term, so for instance the first partial sum is:

$$1-\frac{1}{2},$$

the second partial sum is

$$\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right) = 1-\frac{1}{3},$$

the third partial sum is

$$\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right) = 1-\frac{1}{4},$$

and so on. In general, this pattern where intermediate terms cancel out carries through in all partial sums, so the k-th partial sum is

$$\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\cdots+\left(\frac{1}{k-1}-\frac{1}{k}\right)+\left(\frac{1}{k}-\frac{1}{k+1}\right) = 1-\frac{1}{k+1}.$$

A series converges *by definition* when its sequence of partial sums converges, so since the sequence of partial sums in this case is

$$1-\frac{1}{k+1},$$

which converges to 1 as k goes to infinity, our given series converges to 1.

Careful. The infinite sum we considered above looks like

$$\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\cdots.$$

If we regroup terms like so:

$$1+\left(-\frac{1}{2}+\frac{1}{2}\right)+\left(-\frac{1}{3}+\frac{1}{3}\right)+\left(-\frac{1}{4}-\frac{1}{4}\right)+\cdots,$$

it might at first glance make sense to say that this is

$$1+0+0+0+\cdots$$

since each term the parentheses in the regrouped expression is 0. This seems to suggest that the series should converge and indeed have the value 1. However, we have to be careful with this type of reason. For instance, instead consider the series

$$\sum_{n=0}^{\infty}(-1)^n = 1-1+1-1+1-1+\cdots.$$

Grouping terms like this:

$$(1-1)+(1-1)+(1-1)+\cdots$$

would suggest the value is $0+0+0+0+\cdots=0$, while grouping terms like this:

$$1+(-1+1)+(-1+1)+(-1+1)+\cdots$$

suggests the value is $1+0+0+0+\cdots=1$. Of course, the series $\sum_{n=0}^{\infty}(-1)^n$ is actually divergent (i.e. not convergent), so neither of these "values" is valid. The point is that trying to manipulate an infinite sum (by regrouping terms) as if it were a finite sum can lead to trouble.

If you go on YouTube or various other places you might also come across a "proof" that

$$\sum_{n=1}^{\infty} n = 1+2+3+4+\cdots = -\frac{1}{12}.$$

This is nonsense: the series in question diverges, and so it makes no sense to manipulate it (in the way these videos go through to end up with a value of $-\frac{1}{12}$ as if it where a finite sum. Such manipulations are just "cute" and not actual mathematics.

Convergence tests. The main question we will be interested in for now is determining whether or not a series converges. So far we can only easily answer this for geometric series, or possibly simple examples of telescoping series. But most series are not of these types, so we need more convergence tests. This is what we'll focus on over the next few lectures; in particular we'll look at the Integral Test, the Comparison Test, the Limit Comparison Test, the Alternating Series Test, and the Ratio Test. It is crucial to learn to recognize the types of series each of these tests is best suited for. Our eventual goal is to understand how we can represent functions as series, but the point is that first we need to better understand when series actually converge.

Integral test. The integral test applies to series of the form

$$\sum_{n=1}^{\infty} f(n)$$

where f is a continuous, positive, *decreasing* function. The integral test says that in this setting, the series $\sum_{n=1}^{\infty} f(n)$ and the improper integral $\int_1^\infty f(x)\,dx$ both behave in the same way, meaning they both converge or they both diverge. This is useful since it is usually simpler to determine whether or not an improper integral converges since we can often try to compute its actual value. So, we can turn problems about series convergence into ones about integral convergence instead. Note that the integral test does NOT say that

$$\sum_{n=1}^{\infty} f(n) \quad \text{and} \quad \int_1^\infty f(x)\,dx$$

have the same value, only that both values are either finite or infinite at the same time. We'll give some intuition behind the integral test after the first example below.

Example 1. Consider the series

$$\sum_{n=1}^{\infty} \frac{1}{n}.$$

This is known as the *harmonic* series. The terms $\frac{1}{n}$ in our series come from plugging n into the function

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

Since this function is positive and decreasing, the integral test applies. (Note that this function is decreasing since increasing x makes the value $f(x)$ smaller. Another way to see this is that the derivative $f'(x) = -\frac{1}{x^2}$ is always negative.) So, we instead consider the integral

$$\int_1^\infty \frac{1}{x}\,dx,$$

which we can work with directly. We get:

$$\int_1^\infty \frac{1}{x}\,dx = \lim_{b\to\infty} \int_1^b \frac{1}{x}\,dx = \lim_{b\to\infty} \ln x \Big|_1^b = \lim_{b\to\infty} (\ln b - 1) = \infty.$$

Thus $\int_1^\infty f(x)\,dx$ diverges, so the integral test says that the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges as well.

Note here that even though the terms of this series $\frac{1}{n}$ approach zero, the *series* obtained by adding these terms together does NOT converge. We've mentioned previously that if $\lim_{n\to\infty} a_n$ is not zero, then $\sum a_n$ for sure diverges, and this now an examples where $\lim_{n\to\infty} a_n = 0$ and yet $\sum a_n$ diverges. In general, knowing that $\lim_{n\to\infty} a_n = 0$ says *nothing* about whether or not $\sum a_n$ convergence, so a different convergence test is needed. The point of this example is that even though the terms being added on at each step in

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots$$

are getting smaller and smaller, the sum of *all* of them is actually infinite. (Surprising, no?)

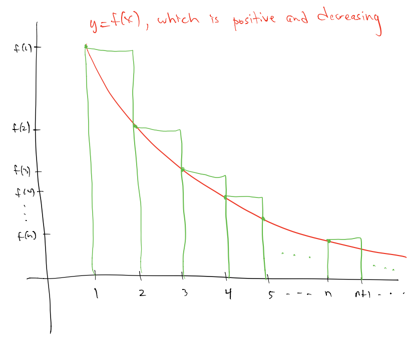
Why does the integral test work? To get a glimpse as to why the integral test works, we note that the series

$$\sum_{n=1}^{\infty} f(n)$$

in the setup of the integral test can be interpreted as a sum of areas. Indeed, think of each $f(n)$ term as

$$f(n) \cdot 1,$$

which is the area of a rectangle of height $f(n)$ and base length 1. Draw these rectangles like so, where the n-th one has base given by the interval $[n, n+1]$:



The sum of the areas of these (infinitely many) rectangles is:

$$f(1) + f(2) + f(3) + f(4) + \cdots = \sum_{n=1}^{\infty} f(n).$$

The integral test is saying that this sum of areas is finite if and only if the area under the curve $y = f(x)$ from $x = 1$ to infinity is finite.

Based on the picture above, we can see that the sum of the areas of the rectangles is actually larger than the area under the curve, so:

$$\sum_{n=1}^{\infty} f(n) \geq \int_1^{\infty} f(x)\, dx \geq 0.$$

Thus certainly if the series is finite so is the integral. Even though the picture doesn't make this clear, it is also true that if the integral is finite so is the series. The idea is that even though these values aren't the same, they are pretty close to one another so that either both are finite or both are infinite simultaneously, which is the statement of the integral test. Again, this is not a proof, but is meant to suggest that there should be some relation between the value of $\sum_{n=1}^{\infty} f(n)$ and that of $\int_1^{\infty} f(x)\, dx$.

p-series test. The same type of integral computation applies to series of the form

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

for some number p. The function $f(x) = \frac{1}{x^p}$ is positive and decreasing on the interval $[1, \infty)$, so the integral test applies to say that

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converges if and only if } \int_1^{\infty} \frac{1}{x^p}\, dx \text{ converges.}$$

This type of integral is one we considered back when we discussed improper integrals, and the fact was that it converged only when $p > 1$. Thus we get that

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converges if } p > 1, \text{ and diverges if } p \leq 1.$$

Such series are known as *p-series*, and so we'll call this is the p-series test.

Fun fact. Here is an interesting fact, which goes way beyond the scope of this course: the actual value of the p-series above when $p = 2$ is $\frac{\pi^2}{6}$:

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{6}.$$

That is, if you could literally add up all the infinitely many terms on the left, you would get $\frac{\pi^2}{6}$ as the result. This is surprising since it is in no way clear what π has to do with the sum of such fractions, and the π seems to come out of nowhere. You might see a justification of this value if you ever learn about *Fourier series*, but this is not something we'll come back to in this course. So, no, you do NOT have to know that $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$ for the purposes of an exam; this was only meant to illustrate an interesting observation.

Example 2. Finally we consider the series

$$\sum_{n=2}^{\infty} n e^{-n^2}.$$

First, the fact that we are starting at $n = 2$ instead of $n = 1$ is not important: in this case we simply use the integral test with the integral $\int_2^{\infty} x e^{-x^2}\, dx$ instead. The fact that xe^{-x^2} is possible to integrate (using the substitution $u = -x^2$) without too much trouble is what suggests the integral test may be useful To be sure the integral test applies, we consider the function

$$f(x) = xe^{-x^2}.$$

This is positive for $x \geq 2$, and

$$f'(x) = e^{-x^2} - 2x^2 e^{-x^2} = (1 - 2x^2)e^{-x^2}$$

is negative for $x \geq 2$ since $1 - 2x^2$ is negative but e^{-x^2} is positive. Hence the integral test is applicable.

We have

$$\begin{aligned} \int_2^{\infty} xe^{-x^2}\, dx &= \lim_{b \to \infty} \int_2^b xe^{-x^2}\, dx \\ &= \lim_{b \to \infty} -\frac{1}{2} e^{-x^2} \Big|_2^b \end{aligned}$$

$$= \lim_{b\to\infty} -\frac{1}{2}(e^{-b^2} - e^{-4})$$
$$= \frac{1}{2}e^{-4}.$$

(To be clear, to evaluate the integral $\int_2^b xe^{-x^2}\,dx$ we used the substitution $u = -x^2$.) Thus $\int_2^\infty xe^{-x^2}\,dx$ converges, so the integral test says that $\sum_{n=2}^\infty ne^{-n^2}$ converges as well.

Lecture 18: Comparison Test

Warm-Up 1. We determine whether the series

$$\sum_{n=3}^\infty \frac{n^2}{n^3+1}$$

converges or diverges. The fact that $\frac{x^2}{x^3+1}$ is possible to integral using a substitution suggests that the integral test may be useful. (The integral test is not the only thing which works here; we'll see later that we can also use the *limit comparison test*.) First we verify that the integral test is actually applicable. The function

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

is positive for $x \geq 3$, and since

$$f'(x) = \frac{(x^3+1)2x - x^2(3x^2)}{(x^3+1)^2} = \frac{2x - x^4}{(x^3+1)^2} = \frac{2x(1-x^3)}{(x^3+1)^2}$$

is negative for $x \geq 3$ (since the numerator is negative and the denominator positive), the integral test is indeed applicable.

We have:

$$\int_3^\infty \frac{x^2}{x^3+1}\,dx = \lim_{b\to\infty}\int_3^b \frac{x^2}{x^3+1}\,dx$$
$$= \lim_{b\to\infty} \frac{1}{3}\ln|x^3+1|\Big|_3^b$$
$$= \lim_{b\to\infty} \frac{1}{3}[\ln(b^3+1) - \ln 10]$$
$$= \infty,$$

where the integral $\int_3^b \frac{x^2}{x^3+1}\,dx$ was computed using the substitution $u = x^3+1$. Thus $\int_3^\infty \frac{x^2}{x^3+1}\,dx$ diverges, so the integral says that $\sum_{n=3}^\infty \frac{n^2}{n^3+1}$ diverges as well.

Warm-Up 2. We determine whether the series

$$\sum_{n=2}^\infty \frac{1}{n\ln n}$$

converges or diverges. The function

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

is positive for $x \geq 2$, and

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

is negative for $x \geq 2$, so f is decreasing. Hence the integral test is applicable.

We have:

$$\begin{aligned} \int_2^\infty \frac{1}{x \ln x}\, dx &= \lim_{b \to \infty} \int_2^b \frac{1}{x \ln x}\, dx \\ &= \lim_{b \to \infty} \ln |\ln x| \Big|_2^b \\ &= \lim_{b \to \infty} (\ln \ln b - \ln \ln 2) \\ &= \infty, \end{aligned}$$

so $\int_2^\infty \frac{1}{x \ln x}\, dx$ diverges. Hence $\sum_{n=2}^\infty \frac{1}{n \ln n}$ diverges by the integral test. (The integral of $\frac{1}{x \ln x}$ was computed using the substitution $u = \ln x$.)

If instead we considered the series

$$\sum_{n=2}^\infty \frac{1}{n(\ln n)^2},$$

the same technique would apply. The function $g(x) = \frac{1}{x(\ln x)^2}$ is still positive and decreasing (which can be checked by seeing that the derivative is negative), so the integral test applies. In this case we have:

$$\int_2^\infty \frac{1}{x(\ln x)^2}\, dx = \lim_{b \to \infty} \int_2^b \frac{1}{x(\ln x)^2}\, dx = \lim_{b \to \infty} -\frac{1}{\ln x}\Big|_2^b = \lim_{b \to \infty} -\left(\frac{1}{\ln b} - \frac{1}{\ln 2}\right) = \frac{1}{\ln 2}.$$

Hence $\int_2^\infty \frac{1}{x(\ln x)^2}\, dx$ converges, so $\sum_{n=2}^\infty \frac{1}{n(\ln n)^2}$ converges as well by the integral test.

Comparison test. The comparison test (and the limit comparison test we'll look at next time) allows us to determine whether or not a series converges by *comparing* to a simpler series whose convergence/divergence is simpler to determine. This test applies to series consisting of all positive terms, so to something like

$$\sum_{n=0}^\infty a_n \text{ where each } a_n \text{ is positive.}$$

The key facts to remember are:

> if the larger series converges, so does the smaller one; and, if the smaller series diverges, so does the larger one.

The point is that such a series of positive terms definitely cannot result in a negative value, so the value is either some positive number or is infinite. So, the only thing we need to determine is whether or not the series has an infinite value (so diverges) or a finite value (so converges).

The first thing we need to do with such series is come up with a *guess* as to whether it should converge or diverge, so that we know which type of comparison we'll need to apply. This is also important because the way in which we come up with this guess will often suggest with which series we should compare our given one. Let's work out some examples to see how this works.

Example 1. Consider the series

$$\sum_{n=1}^\infty \frac{10n^2 - 3n - 1}{n^4 + n^2 + 1}.$$

First we need a guess. The things to focus on are the "dominant" terms in the sequence

$$\frac{10n^2 - 3n - 1}{n^4 + n^2 + 1}.$$

In this case, the dominant term (i.e. the term that overpowers everything else) in the numerator is $10n^2$, while the dominant term in the denominator is n^4. This suggests that, roughly, our given series should behave in a "similar" way to the series

$$\sum_{n=1}^{\infty} \frac{10n^2}{n^4} = \sum_{n=1}^{\infty} \frac{10}{n^2}.$$

This latter series converges by the p-series test, so we make an educated guess that our given series does as well.

Now to make this guess actually precise, we can make a comparison. Note that

$$\frac{10n^2 - 3n - 1}{n^4 + n^2 + 1} \leq \frac{10n^2}{n^4} = \frac{10}{n^2},$$

where the inequality came from making the numerator bigger and the denominator smaller. This equality implies that

$$0 \leq \sum_{n=1}^{\infty} \frac{10n^2 - 3n - 1}{n^4 + n^2 + 1} \leq \sum_{n=1}^{\infty} \frac{10}{n^2}.$$

Since the sum on the right is finite (since $\sum_{n=1}^{\infty} \frac{10}{n^2}$ converges), the first infinite sum should be finite as well. (This is what it means to say that "if the larger series converges, so does the smaller one".) Thus

$$\sum_{n=1}^{\infty} \frac{10n^2 - 3n - 1}{n^4 + n^2 + 1}$$

converges by the comparison test. Note that the series $\sum \frac{10}{n^2}$ we compared our given one with wasn't just pulled out of thin air, but rather came from the series we used in our guess.

Example 2. Consider now the series

$$\sum_{n=2}^{\infty} \frac{10n^4 + n^2 + n + 1}{n^5 - n^4 - 3}.$$

The dominant term in the numerator is $10n^4$ and in the denominator it is n^5. Thus our series should behave in a manner roughly similar to

$$\sum_{n=2}^{\infty} \frac{10n^4}{n^5} = \sum_{n=2}^{\infty} \frac{10}{n},$$

which diverges since it is just 10 times the divergent series $\sum_{n=2}^{\infty} \frac{10}{n}$. So, we guess that our given series diverges.

To show that it diverges using the comparison test, we have to find a "smaller' series which diverges. (In the first example we wanted to show that the given series converged, which meant we had to compare it with a *larger* series which converged.) In this case we have:

$$\frac{10n^4}{n^5} \leq \frac{10n^4 + n^2 + n + 1}{n^5 - n^4 - 3}$$

since the fraction on the left has a smaller numerator and larger denominator than the one on the right. This implies that

$$0 \leq \sum_{n=2} \frac{10n^4}{n^5} \leq \sum_{n=2}^{\infty} \frac{10n^4 + n^2 + n + 1}{n^5 - n^4 - 3},$$

so since the smaller series $\sum_{n=2}^{\infty} \frac{10}{n}$ diverges, so does the larger one. Hence

$$\sum_{n=2}^{\infty} \frac{10n^4 + n^2 + n + 1}{n^5 - n^4 - 3}$$

diverges by the comparison test. Again, note that series we used to compare our given one to came from our educated guess.

Lecture 19: Limit Comparisons and Alternating Series

Warm-Up 1. Consider the series

$$\sum_{n=5}^{\infty} \frac{e^{-n}}{n^2 + 3}.$$

The e^{-n} term is getting smaller and smaller as n increases, so, if we think of this term as $1 \cdot e^{-n}$, in a sense the "dominant" term in the numerator is 1. In the denominator the dominant term is n^2, so the series should behave similarly to

$$\sum_{n=5}^{\infty} \frac{1}{n^2},$$

which converges by the p-series test. Thus we guess that our original series converges.

Since we want to show the original series converges, if we want to apply the comparison test we need to come up with a larger series which converges. We have:

$$\frac{e^{-n}}{n^2 + 3} \leq \frac{1}{n^2},$$

which we can see is true by noting that the numerator on the right is larger than the one on the right and the denominator is smaller. Thus since the larger series $\sum_{n=5}^{\infty} \frac{1}{n^2}$ converges, so does the smaller one, so

$$\sum_{n=5}^{\infty} \frac{e^{-n}}{n^2 + 3}$$

converges by the comparison test.

Warm-Up 2. Consider the series

$$\sum_{n=10}^{\infty} \frac{\ln n}{\sqrt{n} - 1}.$$

Note here that

$$\frac{\ln n}{\sqrt{n} - 1} \geq \frac{1}{\sqrt{n}},$$

since $\ln n > 1$ for $n \geq 10$ and $\sqrt{n} - 1 < \sqrt{n}$. Since the smaller series

$$\sum_{n=10}^{\infty} \frac{1}{\sqrt{n}} = \sum_{n=10}^{\infty} \frac{1}{n^{1/2}}$$

diverges (using the p-series test with $p = 1/2 < 1$), the larger series

$$\sum_{n=10}^{\infty} \frac{\ln n}{\sqrt{n} - 1}$$

diverges as well.

In this example finding the right series to compare the given one too is a little tricker; the "dominant" term in the numerator is $\ln n$, so going by the previous examples we've seen we might think of using

$$\sum_{n=5} \frac{\ln n}{\sqrt{n}}$$

as the comparison series instead. However, it is no simpler to see that this series diverges as opposed to the original one, so this comparison doesn't lead to anywhere. The point is that the idea of "focusing on dominant terms" is a good first attempt which often works, but not always. As you see more and more examples it should become easier to pick up on which comparisons to use when.

Example where comparison doesn't quite work. Consider the series

$$\sum_{n=1}^{\infty} \frac{n^2}{n^3 + n + 1}.$$

Focusing on dominant terms suggests that this series should diverge since

$$\sum_{n=1}^{\infty} \frac{n^2}{n^3} = \sum_{n=1}^{\infty} \frac{1}{n}$$

diverges. If we wanted to use the comparison test to show that our original series diverged, we would have to find a smaller series which diverged. The simplest inequality we can use given the terms of our series is

$$\frac{n^2}{n^3 + n + 1} \leq \frac{n^2}{n^3} = \frac{1}{n},$$

which comes from making the denominator smaller. However, this does us no good: in this case, the divergent series $\sum \frac{1}{n}$ is the *larger* one, and knowing that the larger series diverges tells us nothing about the smaller one. So, doing an ordinary comparison between our original series and $\sum \frac{1}{n}$ leads us nowhere.

However, the guess that our given series should diverge since it should be similarly to $\sum \frac{1}{n}$ was a good one, we just need another way to make this precise. Here is where the *limit comparison test* comes in; this is also a way to compare a given series with another, but where we don't have to worry about which series is "larger" and which is "smaller". To match up with the notation we'll use in a second when stating the limit comparison test in general, denote the terms of our given series and the one we're comparing it to by:

$$a_n = \frac{n^2}{n^3 + n + 1} \quad \text{and} \quad b_n = \frac{1}{n}.$$

We compute the limit $\lim_{n\to\infty} \frac{a_n}{b_n}$:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\frac{n^2}{n^3+n+1}}{\frac{1}{n}} = \lim_{n\to\infty} \frac{n^3}{n^3+n+1} = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n^2}+\frac{1}{n^3}} = 1,$$

where the third equality came from multiplying numerator and denominator by $\frac{1}{n^3}$. Since this limit exists and is positive, the limit comparison test says that

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \frac{n^2}{n^3+n+1} \text{ indeed diverges since } \sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} \frac{1}{n} \text{ diverges.}$$

Limit comparison test. Here is the statement. First, as with the comparison and integral tests, the limit comparison test only applies to series consisting of positive terms. For such series $\sum a_n$ and $\sum b_n$, look at the limit

$$L = \lim_{n\to\infty} \frac{a_n}{b_n}.$$

The *limit comparison test* says that if this limit exists and is positive, then both series $\sum a_n$ and $\sum b_n$ behave in the same way, meaning they both converge or they both diverge. Compared to the ordnary comparison test, we still need to come up with a series to which we can compare our given one, but the benefit is that now we don't have to work with inequalities.

Another example. Consider the series

$$\sum_{n=5}^{\infty} \frac{n(1+e^{-n})}{n^3+3}.$$

All of the terms in this series are positive, and focusing on dominant terms suggests that this series should behave similarly to

$$\sum_{n=5}^{\infty} \frac{n}{n^3} = \sum_{n=5}^{\infty} \frac{1}{n^2},$$

which converges. Thus we guess that our given series should converge too.

To show this, we use the limit comparison test with the series $\sum \frac{1}{n^2}$ we used in our guess. We look at the limit:

$$\lim_{n\to\infty} \frac{\frac{n(1+e^{-n})}{n^3+3}}{\frac{1}{n^2}},$$

which is $\lim_{n\to\infty} \frac{a_n}{b_n}$ in the case where a_n denote the terms of our given series and b_n the terms of the one we are comparing it to. After simplifying, we get:

$$\lim_{n\to\infty} \frac{\frac{n(1+e^{-n})}{n^3+3}}{\frac{1}{n^2}} = \lim_{n\to\infty} \frac{n^3(1+e^{-n})}{n^3+3} = \lim_{n\to\infty} \frac{1+e^{-n}}{1+\frac{3}{n^3}} = 1.$$

Thus $\lim_{n\to\infty} \frac{a_n}{b_n}$ exists (meaning it is finite) and is positive, so since $\sum \frac{1}{n^2}$ converges,

$$\sum_{n=5}^{\infty} \frac{n(1+e^{-n})}{n^3+3}$$

converges as well by the limit comparison test.

Intuition behind limit comparison. This is something you crucially have to know, but it's nice to get a sense for why the limit comparison test works. The idea is that we can think of

$$L = \lim_{n\to\infty} \frac{a_n}{b_n} \text{ existing and being positive}$$

as saying that $\frac{a_n}{b_n}$ should get closer and closer to L as n gets larger, so that (after multiplying through by b_n)

$$a_n \text{ itself gets closer and closer to behaving like } Lb_n.$$

This suggests that

$$\sum a_n \text{ should behave similar to how } \sum Lb_n = L\sum b_n \text{ behaves,}$$

so either both $\sum a_n$ and $\sum b_n$ give finite values (i.e. converge), or both give infinite values (i.e. diverge). Again, this is not an actual proof, just some intuition.

Careful. When $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$, the limit comparison test gives us no information. For instance, consider the series

$$\sum_{n=5}^{\infty} \frac{e^{-n}}{n^2+3}$$

we saw in the first Warm-Up. If we try to use a limit comparison with the series $\sum_{n=5}^{\infty} \frac{1}{n^2}$, we would get:

$$\lim_{n\to\infty} \frac{\frac{e^{-n}}{n^2+3}}{\frac{1}{n^2}} = \lim_{n\to\infty} \frac{n^2 e^{-n}}{n^2+3} = \lim_{n\to\infty} \frac{e^{-n}}{1+\frac{3}{n^2}} = 0.$$

Since we got a limit of zero for $\lim_{n\to\infty} \frac{a_n}{b_n}$, the limit comparison test does not apply. However, in this case, as we saw in the Warm-Up, the ordinary comparison test does apply.

This suggests that there is not always one single test tor try: sometimes one test works when another doesn't, and sometimes multiple tests work. As I said last time, getting used to recognizing which type of test to use in which scenarios is something which comes with much practice.

Alternating series. An *alternating series* is a series where the terms alternate between being positive and negative, or negative and positive. For example,

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \cdots$$

is an alternating series. There is nice, simple test for convergence of such series: the *alternating series test.*

In general, an alternating series can be written as

$$\sum_{n=1}^{\infty} (-1)^n b_n \text{ where the } b_n \text{ are positive.}$$

In other words, b_n is what you get when you factor out -1 from the negative terms and keep the positive terms as they are. In the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = \sum_{n=1}^{\infty} (-1)^n \frac{1}{n},$$

we have $b_n = \frac{1}{n}$. (The alternating series test also applies to something like $\sum_{n=1}^{\infty}(-1)^{n-1}b_n$, meaning that whether we have $(-1)^n$ or $(-1)^{n-1}$ or something else like $(-1)^{n+4}$ is not important; all that matters is that we have signs which alternate between positive/negative or negative/positive.) The alternating series test says that:

if b_n is decreasing and $\lim_{n\to\infty} b_n = 0$, then $\sum_{n=1}^{\infty}(-1)^n b_n$ converges.

Thus, for an alternating series, we can demonstrate convergence simply by showing that the terms of the series approach 0 and (after we forget any negative signs) are decreasing, which means that each term is smaller than the one which came before.

Example. We apply the alternating series test to the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}.$$

As we said above, in this case we consider $b_n = \frac{1}{n}$. These terms are definitely decreasing since the denominators get larger as n increases (we can also see they decreasing by showing that $f(x) = \frac{1}{x}$ has negative derivative for $x \geq 1$) and since

$$\lim_{n\to\infty} \frac{1}{n} = 0,$$

the alternating series test shows that our given series converges.

Fact. The actual value of the series above is:

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln 2.$$

This is not important for the midterm, but is actually something we'll be able to determine once we talk about *power series*. Nonetheless, without knowing that the value is actually $-\ln 2$, we can ask if there is a way in which we can *approximate* the correct value. We'll touch on this a bit next time, and this will play a larger role towards the end of the quarter when we talk about using power series to approximate functions.

Lecture 20: Ratio Test and Absolute Convergence

Warm-Up. We show that the alternating series

$$\sum_{n=2}^{\infty} \frac{(-1)^n n}{n^2+1}$$

converges. In the notation of the alternating series test, here we have

$$b_n = \frac{n}{n^2+1}.$$

First, the function $f(x) = \frac{x}{x^2+1}$ is decreasing since its derivative

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

is negative for $x \geq 2$. Thus $b_n = \frac{n}{n^2+1}$ is decreasing. Next,

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} \frac{n}{n^2+1} = \lim_{n\to\infty} \frac{\frac{1}{n}}{1+\frac{1}{n^2}} = 0.$$

Thus the alternating series test says that $\sum_{n=2}^{\infty} \frac{(-1)^n n}{n^2+1}$ does converge.

Approximating values. Alternating series give us our first example of a series where it is possible to come up with good approximations to the value of a series in cases where we can't determine the actual value. Say that $\sum_{n=1}^{\infty}(-1)^n b_n$ denotes an alternating series. Recall that k-th partial sum of this series is the sum of the first however many terms up to $n = k$:

$$-b_1 + b_2 - b_3 + \cdots + (-1)^k b_k.$$

The idea is that as we get partial sums with more and more terms, such partial sums should be giving better and better approximations to the actual value of the series in question. In particular, the expression

$$|(\text{actual value}) - (k\text{-th partial sum})|,$$

so absolute value of the difference between the actual value and a partial sum approximation, is precisely telling us how good of an approximation the k-th partial sum is to the actual value. Our goal is to be able to "control" how bad this "error' term can be.

For an alternating series, the fact is that this error term can be bounded by one of the b_n terms itself. To be clear, the fact is that

$$\left|(\text{actual value of } \sum(-1)^n b_n) - (k\text{-th partial sum})\right| \leq b_{k+1}.$$

So, the point is that the "error" in approximating the actual value of $\sum(-1)^n b_n$ with the partial sum

$$-b_1 + b_2 - b_3 + \cdots + (-1)^k b_k$$

is no more than the value of b_{k+1}. The smaller this "error" is, the better an approximation we have. The intuition for this fact comes from the following. The actual value of an alternating series is an infnite sum

$$b_0 - b_1 + b_2 - b_3 + \cdots - b_k + b_{k+1} - b_{k+2} + \cdots$$

and the k-th partial sum is something like

$$b_0 - b_1 + b_2 - b_3 + \cdots - b_k.$$

The difference between these is the portion of the infinite sum starting with the b_{k+1} term:

$$b_{k+1} - b_{k+2} + \cdots,$$

and these difference is smaller than b_{k+1} itself since this difference is b_{k+1} minus something smaller, plus something smaller still, minus something even smaller, and so on. (Here we use the fact that the b_n are positive and decreasing towards zero.)

As mentioned previously, having such a bound on the "error" gives a way to determine how good our approximations actually are. This is tough to come up with for series in general, although we'll see a way to proceed with *Taylor series* later on, where this idea will truly shine.

Example. Going back to a series we saw last time:

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n},$$

in this case we get that the (absolute value of the) difference between the actual value (which we claimed last time was $-\ln 2$) and the value of the k-th partial sum is bounded by:

$$|(\text{actual value}) - (k\text{-th partial sum})| \leq \frac{1}{k+1}.$$

For instance, looking at the 9-th partial sum gives:

$$\left|(\text{actual value}) - \left(-1 + \frac{1}{2} - \frac{1}{3} + \cdots + \frac{1}{8} - \frac{1}{9}\right)\right| \leq \frac{1}{10}.$$

So, the error between the actual value of $-\ln 2$ and the approximation given by

$$-1 + \frac{1}{2} - \frac{1}{3} + \cdots - \frac{1}{9}$$

is no more than $\frac{1}{10} = 0.1$. This partial sum is (if you work it out) roughly -0.7456, so the error between this and the actual value is

$$|-\ln 2 - (-0.7456)| \approx 0.005,$$

which is indeed less than 0.1.

Such a bound on the error allows us to answer this type of question: which partial sum would we need to take in order to ensure that the approximated value agrees with the actual value to at least two decimal places? The point is that we know

$$|(\text{actual value}) - (k\text{-th partial sum})| \leq \frac{1}{k+1},$$

so all we need is to guarantee that $\frac{1}{k+1}$ is small enough to force that the actual value and approximate value are within $\frac{1}{1000}$ of one another. Indeed, two numbers that are within $\frac{1}{1000}$ of one another will necessarily have the same first two decimal places. So, in this example, we need a value of k satisfying

$$\frac{1}{k+1} \leq \frac{1}{1000},$$

and for instance $k = 1000$ works. (Of course, $k = 999$ also works, but for now we are not necessarily looking for the most "efficient" value of k.) Thus the 1000-th partial sum of this series would approximate the actual value to least two decimal places. We'll return to such approximation questions later after we talk about Taylor series.

Ratio test. Our final convergence test is the so-called *ratio test*, which is possibly the simplest one to apply when it works. Say we have a series $\sum_{n=1}^{\infty} a_n$. We look at the limit:

$$L = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$$

obtained from the absolute value of the fraction of the $(n+1)$-st term in our series and the n-term. The ratio test says that:

- if $L < 1$, then $\sum a_n$ converges (actually, it converges *absolutely*, which is better; we'll define this term in a bit)
- if $L > 1$ or $L = \infty$, then $\sum a_n$ diverges, and
- if $L = 1$, the ratio test gives no information.

Thus, computing the required limit will tells us whether or series converges or diverges, as long as the limit is not equal to 1.

Example 1. Consider the series
$$\sum_{n=1}^{\infty} (-1)^n \frac{n}{5^n}.$$
In the notation of the ratio test, $a_n = (-1)^n \frac{n}{5^n}$, so we compute the limit:
$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{|(-1)^{n+1}\frac{n+1}{5^{n+1}}|}{|(-1)^n \frac{n}{5^n}|} = \lim_{n\to\infty} \frac{(n+1)5^n}{n5^{n+1}} = \lim_{n\to\infty} \frac{(n+1)}{5n} = \frac{1}{5}.$$
Note that the $(-1)^n$ terms disappeared after taking absolute values. Thus since this limit is less than 1, the ratio test tells us that this series converges. (Actually, it tells us that it converges absolutely, which we still have yet to define.)

Example 2. Next we consider
$$\sum_{n=0}^{\infty} \frac{(-3)^n}{(2n+1)!}.$$
In the ratio test we need the following limit:
$$\lim_{n\to\infty} \frac{|\frac{(-3)^{n+1}}{(2(n+1)+1)!}|}{|\frac{(-3)^n}{(2n+1)!}|} = \lim_{n\to\infty} \frac{3^{n+1}}{(2n+3)!} \frac{(2n+1)!}{3^n} = \lim_{n\to\infty} \frac{3}{(2n+3)(2n+2)} = 0.$$
To be clear, here we used the fact that
$$(2n+3)! = (2n+3)(2n+2)(2n+1)(2n)(2n-1)\cdots 2\cdot 1 \text{ and } (2n+1)! = (2n+1)(2n)(2n-1)\cdots 2\cdot 1$$
in order to simplify
$$\frac{(2n+1)!}{(2n+3)!} = \frac{1}{(2n+3)(2n+2)}.$$
Since we got a limit of zero, the ratio test implies that our given series converges.

Example 3. For
$$\sum_{n=1}^{\infty} n2^n,$$
the ratio test gives
$$\lim_{n\to\infty} \frac{(n+1)2^{n+1}}{n2^n} = \lim_{n\to\infty} \frac{2(n+1)}{n} = 2.$$
Since this is larger than 1, the given series diverges.

Absolute convergence. As stated above, the actual conclusion of the ratio test in the $L < 1$ case is that the given series converges *absolutely*. To say that $\sum a_n$ is *absolutely convergent* means that the series obtained by taking absolute values $\sum |a_n|$ converges. This always implies that the original series converges as well, so absolute convergence is a special type of convergence. If a series converges but does not converge absolutely, we say it is *conditionally* convergent. For instance, the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

converges by the alternating series test, but the series of absolute values

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges. Thus $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ is converges conditionally.

Next time we'll mention why people care about this distinction between absolute vs conditional convergence. This won't play a big role in this course, so for us the key takeaway is that absolute convergence implies ordinary convergence, which is really all we care about.

Lecture 21: Power Series

Warm-Up 1. We determine whether or not

$$\sum_{n=1}^{\infty} \frac{n!}{100^n}$$

converges. We compute:

$$\lim_{n\to\infty} \frac{\frac{(n+1)!}{100^{n+1}}}{\frac{n!}{100^n}} = \lim_{n\to\infty} \frac{(n+1)!100^n}{100^{n+1}n!} = \lim_{n\to\infty} \frac{n+1}{100} = \infty,$$

where we use the fact that

$$\frac{(n+1)!}{n!} = \frac{(n+1)n!}{n!} = n+1.$$

Since this limit is ∞ (which we consider to be in the $L > 1$ case), the ratio test tells us that this series diverges.

Warm-Up 2. Now consider the series

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{n^2+4}.$$

Applying the ratio test gives:

$$\lim_{n\to\infty} \frac{|(-1)^{(n+1)-1} \frac{n+1}{(n+1)^2+4}|}{|(-1)^n \frac{n}{n^2+4}|} = \lim_{n\to\infty} \frac{(n+1)(n^2+4)}{n(n+1)^2+4} = \lim_{n\to\infty} \frac{n^3+n^2+4n+4}{n^3+2n^2+n+4} = 1.$$

Since we got a limit of 1, the ratio test gives us no information.

However, we can instead show that this series converges using the alternating series test; I'll leave it to you to work out the details. Since it does converge, we can ask whether it converges absolutely. For this we consider the series obtained by taking absolute values:

$$\sum_{n=1}^{\infty} \frac{n}{n^2+4}.$$

You can show that this diverges by comparing it to $\sum_{n=1}^{\infty} \frac{1}{n}$ using either the comparison or limit comparison test. So, since our original series converges but the series of absolute values does not, our original series

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{n^2+4}$$

is conditionally convergent.

Absolute vs conditional. Now we give the reason why the distinction between absolute vs conditional convergence matters. This is NOT something we'll focus on in this course and is more of a "fun fact" everyone should hear about once in their lifetimes. Here is the basic idea: rearranging the terms of an absolutely convergent series does not affect the convergence nor the value, whereas rearranging the terms of a conditionally convergent series *could* affect the value!

If we have an infinite sum:

$$a_1 + a_2 + a_3 + a_4 + \cdots,$$

we can rearrange the terms in some way:

$$a_{100} + a_3 + a_{10} + a_4 + a_1 + a_{567} + \cdots \text{ and so on.}$$

The fact is that new sum does NOT necessarily have the same value as the original one EXCEPT when the original series was absolutely convergent. So, for instance, we've seen previously that

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \cdots$$

is conditionally convergent. Because of this, rearranging terms could definitely have an affect the actual value of the series. Even worse, I claim that there is a rearrangement which gives the value π as a result, there is another rearrangement which gives the value e, another giving the value $2^{\sin 1}$, etc: given any real number whatsoever, there is a way to rearrange the terms of a conditionally convergent series to obtain a series whose value is that chosen real number! This cannot happen for absolutely convergent series, where rearrangements affect nothing.

These facts might seem counterintuitive, since rearranging the terms of a finite sum never affects the value:

$$x + y + z + w \text{ is the same as } x + w + z + y \text{ is the same as } z + w + x + y$$

and so on. This is yet another subtle distinction between infinite and finite sums which shows that we have to be careful applying whatever intuition we have for finite sums to infinite sums. Understanding why infinite sums have the properties listed above is way beyond the scope of this course, and, as I said, this will not play a further role for us. But, it is an interesting observation nonetheless!

Power series. We now move onto the final topic of the quarter, which revolves around the idea of representing *functions* as series. Indeed, this is the whole reason why we ever cared about series in the first place. The point is that representing a function as a series gives us new ways of studying that function, which is especially useful for functions which are otherwise difficult to understand.

The key notion is that of a *power series*, which is a type series involving powers of a variable x. To be clear, a power series is a series of the form

$$\sum_{n=0}^{\infty} c_n(x-a)^n,$$

where the c_n are numbers, x is a variable, and a is a number we call the "center" of the series. (Alternatively we say that this is a power series *centered* at a.) The idea is that, because x is a variable, we view a power series as defining a function depending on x, and that the functions defined by power series often turn out to be functions we all know and love.

The first question to handle is that of convergence. Namely, since x is a variable, we may have that for certain values of x a given power series converges while for other values it diverges. The basic fact is that there is a certain interval (called the *interval of convergence* on which a power series converges—meaning that it converges for values of x in that interval. The radius (i.e half the total length) of this interval is called the *radius of convergence* of the power series.

Example 1. The first basic power series you should have engrained in your minds is:

$$\sum_{n=0}^{\infty} x^n.$$

This power series is centered at 0, since $x^n = (x-0)^n$. Based on what we saw previously when looking at geometric series, we know that this series converges only for values of x in the interval $(-1,1)$, or equivalent for x satisfying $|x| < 1$. Thus, this series has radius of convergence 1 and interval of convergence $(-1,1)$. (In general, a series centered at a will converge for values of x satisfying $|x-a| < R$, where R is the radius of convergence, and will have interval of convergence given by $(a-R, a+R)$ and *possibly* including one or both of the endpoints $a-R$ and $a+R$. We'll soon see examples of this.)

In this case, the given series converges to

$$\frac{1}{1-x},$$

again using what we already know about geometric series. Thus, we say that this series *represents* the function $\frac{1}{1-x}$ on the interval $(-1,1)$, meaning that for values of x in this interval, this series gives the same value as this function:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \text{ for } x \text{ in } (-1,1).$$

We refer to this as a *power series representation* of the function $\frac{1}{1-x}$.

Example 2. Consider now the power series

$$\sum_{n=0}^{\infty} 3^n x^n,$$

which is also centered at 0. It turns out that we can determine the radius and interval of convergence of this series directly from the series we had in Example 1. Indeed, recall the following representation from above:

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n \text{ for } y \text{ in } (-1,1).$$

Here I am denoting the power series variable by y instead of x only to make what we're about to do simpler to understand. Note that if we set y equal to $3x$ in this expression we get:

$$\frac{1}{1-3x} = \sum_{n=0}^{\infty}(3x)^n = \sum_{n=0}^{\infty} 3^n x^n,$$

which is precisely the series we're looking at in this example! The point here is that the series

$$\sum_{n=0}^{\infty} 3^n x^n$$

can be obtained from the series

$$\sum_{n=0}^{\infty} y^n$$

via a certain manipulation, in this case replacing y with $3x$.

We know that $\sum_{n=0}^{\infty} y^n$ converges when $|y| < 1$, so, since our new series was obtained by setting $y = 3x$, our new series should converge when $|3x| < 1$. This gives that x should satisfy

$$|x| < \frac{1}{3}, \text{ or equivalently that } x \text{ should be in } \left(-\frac{1}{3}, \frac{1}{3}\right).$$

Hence the series

$$\sum_{n=0}^{\infty} 3^n x^n$$

has radius of convergence $\frac{1}{3}$ and interval of convergence $(-\frac{1}{3}, \frac{1}{3})$. We can say that

$$\frac{1}{1-3x} = \sum_{n=0}^{\infty} 3^n x^n$$

represents the function $\frac{1}{1-3x}$ as a power series centered at 0 on the interval $(-\frac{1}{3}, \frac{1}{3})$.

Example 3. The power series

$$\sum_{n=0}^{\infty}(3x-1)^n$$

can also be derived from the standard

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n.$$

First, to be clear, why is $\sum(3x-1)^n$ a power series? We defined a power series to be one involving powers of $x - a$ for some center a, but here we have powers of $3x - 1$. The point is that we can rewrite this series to make it be of the correct form:

$$\sum_{n=0}^{\infty}(3x-1)^n = \sum_{n=0}^{\infty}\left[3\left(x-\frac{1}{3}\right)\right]^n = \sum_{n=0}^{\infty} 3^n \left(x-\frac{1}{3}\right)^n.$$

(To be clear, we wrote $3x-1$ as $3(x-\frac{1}{3})$.) Thus this is a power series centered at $\frac{1}{3}$.

Now, we can obtain this given series from

$$\frac{1}{1-y}=\sum_{n=0}^{\infty}y^n$$

by setting $y=3x-1$:

$$\frac{1}{1-(3x-1)}=\sum_{n=0}^{\infty}(3x-1)^n.$$

This series will converge when $|y|<1$, so when

$$|3x-1|<1.$$

Taking into account that this should be a power series centered at $\frac{1}{3}$, we write this inequality as

$$3\left|x-\frac{1}{3}\right|<1, \text{ or } \left|x-\frac{1}{3}\right|<\frac{1}{3}.$$

Thus, the given series has radius of convergence $\frac{1}{3}$ and has interval of convergence given by

$$\left(\frac{1}{3}-\frac{1}{3},\frac{1}{3}+\frac{1}{3}\right)=\left(0,\frac{2}{3}\right).$$

(The left endpoint is the center minus the radius, and the right endpoint is the center plus the radius.) Another way of saying this is that since $\sum y^n$ converges for y in $(-1,1)$, our given series should converge for $y=3x-1$ in $(-1,1)$, which says that $3x$ should be in $(0,2)$, and hence that x should be in $(0,\frac{2}{3})$.

We can thus say that

$$\frac{1}{2-3x}=\sum_{n=0}^{\infty}(3x-1)^n=\sum_{n=0}^{\infty}3^n\left(x-\frac{1}{3}\right)^n$$

represents the function $\frac{1}{2-3x}$ as a power series centered at $\frac{1}{3}$ on the interval $(0,\frac{2}{3})$.

Example 4. Of course, not all power series can be derived from a clever manipulation of

$$\frac{1}{1-y}=\sum_{n=0}^{\infty}y^n.$$

The main tool for determining a radius of convergence in general is the *ratio test*. For instance, consider the power series

$$\sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

Setting $a_n=\frac{x^n}{n!}$, the ratio test tells us that this series will converge when

$$\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}<1.$$

In our case we thus consider:

$$\lim_{n\to\infty} \frac{|x|^{n+1}}{(n+1)!} \frac{n!}{|x|^n} = \lim_{n\to\infty} \frac{|x|}{n+1} = 0.$$

Because this is smaller than 1 no matter what x is, we thus conclude that series

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

converges for *all* values of x. The interval of convergence is thus $(-\infty, \infty)$, and we say that the interval of convergence is ∞. (We will see soon enough that this series, amazingly, actually represents the function e^x.)

Example 5. Finally, consider the power series

$$\sum_{n=1}^{\infty} n(x-2)^n.$$

We have:

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{(n+1)|x-2|^{n+1}}{n|x-2|^n} = \lim_{n\to\infty} \frac{(n+1)|x-2|}{n} = |x-2| \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = |x-2|.$$

Thus according to the ratio test, the given series converges when $|x-2| < 1$. Since this is a power series centered at 2, the radius of convergence is thus 1 and the interval of convergence at least includes the interval $(2-1, 2+1) = (1, 3)$.

However, recall that the ratio test is inconclusive when $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = 1$. In our case, this means that the ratio test is inconclusive when $|x-2| = 1$, so when $x = 1$ or $x = 3$. The point is that this method of finding the radius and interval of convergence will say nothing about the endpoints of the resulting interval, and so we have to see what happens at those endpoints separately. When $x = 1$ our given series becomes

$$\sum_{n=1}^{\infty} n(-1)^n,$$

which diverges since $\lim_{n\to\infty} n(-1)^n \neq 0$, and when $x = 3$ our series is

$$\sum_{n=1}^{\infty} n,$$

which also diverges. This means that neither 1 nor 3 are in the interval of convergence, so that the interval of convergence is indeed $(1, 3)$.

In general, it might happen for some power series, one or both of the endpoints might actually be included in the interval of convergence, which we have to check for separately as we did here. For instance, the series

$$\sum_{n=0}^{\infty} \frac{x^n}{n}$$

has radius of convergence 1, which can be determined by the ratio test. This gives an interval of convergence at least consisting of $(-1, 1)$. When $x = 1$ we get the series

$$\sum_{n=0}^{\infty} \frac{1}{n},$$

which diverges, but when $x = -1$ we get

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n},$$

which actually converges by the alternating series test. Thus, for

$$\sum_{n=0}^{\infty} \frac{x^n}{n},$$

-1 should be included in the interval of convergence, so the interval of convergence is the half-open/half-closed interval $[-1, 1)$. Note that we didn't have to check endpoints in the first three examples we looked at since these were derived from the geometric series

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n$$

where we know for sure the interval of convergence $(-1, 1)$ does not include the endpoints.

Lecture 22: Representing Functions as Power Series

Warm-Up 1. We determine the radius and interval of convergence of the power series

$$\sum_{n=2}^{\infty} \frac{(-1)^n (x-2)^n}{2^n (n+1)}.$$

Note that this is a power series centered at 2. We compute:

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{|x-2|^{n+1}}{2^{n+1}(n+2)} \frac{2^n(n+1)}{|x-2|^n} = \lim_{n\to\infty} \frac{|x-2|}{2} \left(\frac{n+1}{n+2}\right) = \lim_{n\to\infty} \frac{|x-2|}{2} \frac{1+\frac{1}{n}}{1+\frac{2}{n}} = \frac{|x-2|}{2}.$$

By the ratio test this series converges when

$$\frac{|x-2|}{2} < 1, \text{ so when } |x-2| < 2.$$

The the radius of convergence is 2 and the interval of convergence is at least $(2-2, 2+2) = (0, 4)$.

Now we check for convergence at the endpoints. For $x = 0$ our series becomes

$$\sum_{n=2}^{\infty} \frac{(-1)^n (-2)^n}{2^n (n+1)} = \sum_{n=2}^{\infty} \frac{1}{n+1},$$

which diverges as we can see doing a limit comparison test with the harmonic series $\sum_{n=2}^{\infty} \frac{1}{n}$. Thus 0 is not in the interval of convergence. For $x = 4$ this series becomes

$$\sum_{n=2}^{\infty} \frac{(-1)^n 2^n}{2^n (n+1)} = \sum_{n=2}^{\infty} \frac{(-1)^n}{n+1},$$

which converges by the alternating series test. (Note that here I'm not working out the details of the alternating series test, but this is something you should be able to do.) Thus 4 is in the interval of convergence, so the interval of convergence of the given series is $(0, 4]$.

Warm-Up 2. We determine the radius and interval of convergence of

$$\sum_{n=0}^{\infty}(2n+1)!(x-1)^n.$$

We have:

$$\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}=\lim_{n\to\infty}\frac{(2(n+1)+1)!|x-1|^{n+1}}{(2n+1)!|x-1|^n}=\lim_{n\to\infty}\frac{(2n+3)!}{(2n+1)!}|x-1|=\lim_{n\to\infty}(2n+3)(2n+2)|x-1|.$$

Now we have to be careful: the $(2n+3)(2n+2)$ portion goes to ∞, but whether we actually get an infinite value for the limit will depend on whether the $|x-1|$ term is zero. Indeed, if $x=1$, $|x-1|=0$ and this limit is:

$$\lim_{n\to\infty}(2n+3)(2n+2)0=\lim_{n\to\infty}0=0,$$

and since this is less than 1 the ratio test tells us that the given series *does* converge when $x=1$. However, if $x\neq 1$, $|x-1|\neq 0$ and so in this case

$$\lim_{n\to\infty}(2n+3)(2n+2)|x-1|=\infty,$$

and hence the series does not converge in this case.

Thus this series only converges when $x=1$, so we say that the interval of convergence (which isn't really an interval in this case) is just the single point 1, and so the radius of convergence is 0. (This makes sense since a single point has zero length, and half of zero is zero.) The various examples we've seen illustrate the different things which can happen: either the radius of convergence is 0 in which case the series converges only at its center, or the radius of convergence is infinite in which case the interval of convergence is $(-\infty,\infty)$, or the radius of convergence is some positive number R in which case the interval of convergence looks like $(a-R,a+R)$ (where a is the center) and possibl includes none, one, or both of the endpoints.

Representing functions as series. We saw last time that we can represent various functions as power series, and now we develop this a bit more. For instance, we can ask: what function does the power series

$$\sum_{n=0}^{\infty}(-1)^n x^{2n}$$

represent? This is something we can determine by manipulating the series

$$\frac{1}{1-y}=\sum_{n=0}^{\infty}y^n$$

along the lines of some examples we saw last time. In this case, setting $y=-x^2$ in this series gives

$$\sum_{n=0}^{\infty}(-x^2)^n=\sum_{n=0}^{\infty}(-1)x^{2n},$$

which is precisely the series we're asking about. Since

$$\frac{1}{1-y}=\sum_{n=0}^{\infty}y^n$$

we thus get that

$$\frac{1}{1-(x^2)} = \sum_{n=0}^{\infty}(-1)^n x^{2n},$$

so the power series we had at the beginning represents the function

$$\frac{1}{1+x^2}.$$

Now, on which interval is this representation valid? That is, for which values of x is it true that

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n}?$$

The answer is given by the interval of convergence of this series. We can find this using the ratio test as in the Warm-Up problems, or we can use the fact that this power series was derived from

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n.$$

We know that this latter series converges only when $|y| < 1$, so the series we obtained should converge only when $|-x^2| < 1$ since we used $y = -x^2$ as our y-value. The inequality $|-x^2| < 1$ simplifies to $|x|^2 < 1$, which in turn simplifies to $|x| < 1$, so we get that

$$\sum_{n=0}^{\infty}(-1)^n x^{2n}$$

has radius of convergence 1 and interval of convergence $(-1, 1)$. Thus the representation

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n}$$

is valid for $-1 < x < 1$.

Example. As another example, again start with

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n \text{ for } |y| < 1.$$

Making the substitution $y = 1 - x$ gives

$$\frac{1}{1-(1-x)} = \sum_{n=0}^{\infty}(1-x)^n \text{ for } |1-x| < 1,$$

which we can write as

$$\frac{1}{x} = \sum_{n=0}^{\infty}(-1)^n (x-1)^n \text{ for } |x-1| < 1.$$

This thus gives a representation of the function $\frac{1}{x}$ as a power series centered at 1, which is valid on the interval $(1-1, 1+1) = (0, 2)$. Of course, this interval of convergence can also be found by using the ratio test to find the radius of convergence, which is 1.

Differentiating and integrating series. Manipulating a power series by making a substitution (such as $y = -x^2$ or $y = 1 - x$) like we did above is one way of manipulating one series in order to produce another. Another way of manipulating a series comes from differentiation and integration.

Write a power series as an infinite sum:

$$\sum_{n=0}^{\infty} c_n(x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots .$$

We can differentiate this just as we're used to, by differentiating each term one at a time:

$$\left(\sum_{n=0}^{\infty} c_n(x-a)^n\right)' = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \cdots .$$

This results in the formula:

$$\left(\sum_{n=0}^{\infty} c_n(x-a)^n\right)' = \sum_{n=0}^{\infty} c_n n(x-a)^{n-1},$$

where the $n(x-a)^{n-1}$ piece comes from differentiating $(x-a)^n$. Note that the $n = 0$ term in the resulting series is 0 (because plugging in $n = 0$ into $c_n n(x-a)^{n-1}$ gives zero), so we can rewrite the series to start at $n = 1$ (the first nonzero term) instead:

$$\left(\sum_{n=0}^{\infty} c_n(x-a)^n\right)' = \sum_{n=1}^{\infty} c_n n(x-a)^{n-1}.$$

This is just reflecting the fact that the constant c_0 term in

$$c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots$$

disappears after taking derivatives. This process is called *term-by-term differentiation* of a power series. We'll see examples of this in a bit.

Similarly, we can integrate one power series in order to produce another. Integrating

$$\sum_{n=0}^{\infty} c_n(x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots$$

should give

$$\int \left(\sum_{n=0}^{\infty} c_n(x-a)^n\right) dx = c_0 x + c_1\frac{(x-a)^2}{2} + c_2\frac{(x-a)^3}{3} + c_3\frac{(x-a)^4}{4} + \cdots ,$$

so

$$\int \left(\sum_{n=0}^{\infty} c_n(x-a)^n\right) dx = \sum_{n=0}^{\infty} c_n\frac{(x-a)^{n+1}}{n+1}.$$

To be clear, the

$$\frac{(x-a)^{n+1}}{n+1}$$

term comes from integrating $(x-a)^n$. When considering indefinite integrals we should also throw on a $+C$ term at the end as usual. This process is called *term-by-term integration.*

Example 1. Consider the series

$$\sum_{n=1}^{\infty}(-1)^n 2nx^{2n-1}.$$

The key observation here is that this series is precisely the result of differentiating the series

$$\sum_{n=0}^{\infty}(-1)^n x^{2n}$$

we saw earlier. Indeed, differentiating term-by-term gives

$$\left(\sum_{n=0}^{\infty}(-1)^n x^{2n}\right)' = \sum_{n=1}^{\infty}(-1)^n 2nx^{2n-1}$$

since $2nx^{2n-1}$ is the derivative of x^{2n}. Since

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n},$$

the series we're looking at in this example should represent the function obtained by differentiating

$$\frac{1}{1+x^2}.$$

In other words, taking derivatives of both sides of

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n}$$

gives

$$\left(\frac{1}{1+x^2}\right)' = \sum_{n=1}^{\infty}(-1)^n 2nx^{2n-1},$$

so the given series in this example represents the function

$$\frac{-2x}{(1+x^2)^2} = \sum_{n=1}^{\infty}(-1)^n 2nx^{2n-1}.$$

Above we phrased this as: given the series

$$\sum_{n=1}^{\infty}(-1)^n 2nx^{2n-1},$$

find the function it represents. But we can also ask this is the other way around: given the function

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

represent it as a series. The key is to recognize that this function is the derivative of

$$\frac{1}{1+x^2}.$$

Since

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n},$$

differentiating both sides gives

$$\frac{-2x}{(1+x^2)^2} = \sum_{n=1}^{\infty}(-1)^n 2nx^{2n-1}$$

as the desired representation.

Example 2. Consider the series

$$\sum_{n=1}^{\infty}\frac{(-1)^n x^n}{n}.$$

Note that if we start with

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n,$$

integrating both sides gives

$$-\ln|1-y| = \sum_{n=0}^{\infty}\frac{y^{n+1}}{n+1} + C.$$

The point is that integration is the type of operation which can give additional n terms (in this case $n+1$) in the denominator of a series expression. The unknown constant of integration C can be found by plugging in $y=0$ into both sides: this gives

$$-\ln 1 = \sum_{n=0}^{\infty} 0 + C, \text{ or } 0 = 0 + C.$$

Hence $C = 0$ so

$$-\ln|1-y| = \sum_{n=0}^{\infty}\frac{y^{n+1}}{n+1}.$$

Making the substitution $y = -x$ then gives

$$-\ln|1+x| = \sum_{n=0}^{\infty}\frac{(-1)^{n+1}x^{n+1}}{n+1}.$$

Now, this is almost the series we want, only that the series we want starts at $n = 1$ and involves

$$\frac{x^n}{n} \text{ instead of } \frac{x^{n+1}}{n+1}.$$

However, note that we can *rewrite* the series

$$\sum_{n=0}^{\infty}\frac{(-1)^{n+1}x^{n+1}}{n+1}$$

to make it start at $n = 1$ instead. To make this clear, let us instead use m as the indexing variable:

$$\sum_{m=0}^{\infty}\frac{(-1)^{m+1}x^{m+1}}{m+1}.$$

If we set $n = m + 1$, we get a series starting at $n = 1$ (since $n = 1$ when $m = 0$) which looks like:

$$\sum_{m=0}^{\infty} \frac{(-1)^{m+1} x^{m+1}}{m+1} = \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n},$$

and this latter series is the one we want. We conclude that

$$-\ln|1+x| = \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n}$$

is the function our given series represents.

As in the previous example, we can phrase this the other way around. Say we want to find the series which represents

$$-\ln|1+x|.$$

We note that this function is obtained by integrating

$$\frac{1}{1+x},$$

so that if we know how to represent this latter function as a series, we can integrate to find a representation of the function we want. Since

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n,$$

which comes from setting $y = -x$ in the standard series expression for $\frac{1}{1-y}$, we get after integrating that:

$$\ln|1+x| = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1} + C.$$

The unknown constant C can be found by setting $x = 0$ in this expression, so $0 = 0 + C$ and hence $C = 0$, and thus

$$\ln|1+x| = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}.$$

We can reindex this series to start at $n = 1$ instead (this will have effect of replacing each n showing up in the series expression with $n - 1$) to get

$$\ln|1+x| = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n},$$

and finally multiplying through by -1 to get

$$-\ln|1+x| = \sum_{n=1}^{\infty} -(-1)^{n-1} \frac{x^n}{n} = \sum_{n=1}^{\infty} (-1)^n \frac{x^n}{n}$$

as a series expression for $-\ln|1+x|$. This series (if you work it out) has interval of convergence $(-1, 1]$, and since 1 is then in this interval, we get the equality

$$-\ln 2 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

after setting $x = 1$. This was a claim I made previously when discussing alternating series, and now we can see why this is the correct value of this series.

Lecture 23: Taylor and Maclaurin Series

Warm-Up 1. We determine the function which is represented by the series

$$\sum_{n=2}^{\infty} n(n-1)x^n.$$

The key observation is that this series is *almost* what we get if we take *two* derivatives of

$$\sum_{n=0}^{\infty} x^n.$$

Indeed, taking one derivative gives

$$\sum_{n=1}^{\infty} nx^{n-1},$$

and taking another gives

$$\sum_{n=2}^{\infty} n(n-1)x^{n-2}.$$

(Note that here we've written this series to start at $n = 2$; we could have written it to start at $n = 0$ instead only that the $n = 0$ term itself would be 0 because of the n coefficient, and the $n = 1$ term would also be zero because of the $n - 1$ coefficient. The first nonzero term is the $n = 2$ term, which is why we write the series to start at this value.) Since the original series $\sum x^n$ represented the function $\frac{1}{1-x}$, the series after we take two derivatives will represent

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

Thus so far we have that

$$\frac{2}{(1-x)^2} = \sum_{n=2}^{\infty} n(n-1)x^{n-2}.$$

The only difference between this series and the one we want is the power of x, but that can be fixed by multiplying through by x^2:

$$\frac{2x^2}{(1-x)^2} = x^2 \sum_{n=2}^{\infty} n(n-1)x^{n-2} = \sum_{n=2}^{\infty} n(n-1)x^2x^{n-2} = \sum_{n=2}^{\infty} n(n-1)x^n.$$

Thus we conclude that the original series at the start of this Warm-Up is a series representation of the function

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

Warm-Up 2. We find a series representation of the function

$$x \ln |1 - x^3|.$$

To build up to this, we start by finding a series representation of

$$\ln |1 - x^3|.$$

Since

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n,$$

integrating both sides gives

$$-\ln|1-y| = \sum_{n=0}^{\infty} \frac{y^{n+1}}{n+1} + C,$$

and plugging in $y = 0$ to get that $C = 0$ leaves us with

$$-\ln|1-y| = \sum_{n=0}^{\infty} \frac{y^{n+1}}{n+1}.$$

We can multiply through by -1 and rewrite the series on the right to start at $n = 1$ (and replacing n by $n-1$ accordingly) to get

$$\ln|1-y| = \sum_{n=1}^{\infty} -\frac{y^n}{n}.$$

Setting $y = x^3$ gives

$$\ln|1-x^3| = \sum_{n=1}^{\infty} -\frac{x^{3n}}{n},$$

where we used the fact that $(x^3)^n = x^{3n}$, and finally multiplying through by x gives

$$x\ln|1-x^3| = \sum_{n=1}^{\infty} -\frac{xx^{3n}}{n} = \sum_{n=1}^{\infty} -\frac{x^{3n+1}}{n}$$

as the desired representation.

Warm-Up 3. Finally we find a series representation for $\arctan x$. The key fact we need is that the derivative of this functions is $\frac{1}{1+x^2}$:

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

Thus if we have a series representation of $\frac{1}{1+x^2}$, integrating term-by-term will give a series representation of $\arctan x$. We've seen before that

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n},$$

which comes from making the substitution $y = -x^2$ in

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n.$$

Integrating both sides of

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}$$

gives

$$\arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} + C.$$

Setting $x = 0$ gives $0 = 0 + C$, so $C = 0$ and thus

$$\arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$$

is the desired representation.

A series which equals its own derivative. Consider the power series

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Differentiating term-by-term gives

$$\left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right)' = \sum_{n=1}^{\infty} \frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!}.$$

Note that we can rewrite this resulting series to start at $n = 0$ by replacing n with $n - 1$ to get

$$\sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

and the upshot is that

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

equals its own derivative!

This suggests that this series should represent a function like e^x, or some function which equals its own derivative. (In fact, the only functions which equal their own derivatives are those of the form ae^x where a is a constant.) And in fact it is true here that

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

The question is: how can we derive this series representation for e^x directly without having to notice that this series equals its own derivative? Or, more generally, how do we derive series representations for other functions, and not just those which can be related to $\frac{1}{1-y}$ in some way? The answer comes from the notion of a *Taylor series*.

Taylor and Maclaurin series. The *Taylor series centered at a* of the function f is the power series which looks like

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n.$$

Here, $f^{(n)}$ denotes the n-th derivative of f, and we interpret $f^{(0)}$ as simply being f itself, undifferentiated. Writing this series out as an infinite sum gives

$$f(a) + f'(a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots,$$

so that the coefficients of a Taylor series come from evaluating a derivative of f at a and then dividing by an appropriate factorial.

The crucial point is that if we are aiming to express a function as a power series centered at a, the *only* series for which this is possible *is* the Taylor series centered at a of that function. We'll come back to this next time to see why, but for now we focus on computing Taylor series directly. A Taylor series centered at 0 is usually called a *Maclaurin series*, so to be clear the Maclaurin series of f is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

Essentially, Maclaurin series are the most basic (and most important) types of Taylor series.

Example 1. We determine the Maclaurin series (i.e. Taylor series centered at 0) of e^x, and then we determine its Taylor series centered at -3. To compute either one we need to start by determining the derivatives of e^x, which is easy in this case since e^x equals its own derivative:

$$f(x) = e^x,\ f'(x) = e^x,\ f''(x) = e^x,\ f^{(3)}(x) = e^x,\ \text{and in general } f^{(n)}(x) = e^x.$$

Evaluating these at 0 gives

$$f^{(n)}(0) = e^0 = 1 \text{ for all } n,$$

so the coefficient of x^n in the Maclaurin series is

$$\frac{f^{(n)}(0)}{n!} = \frac{1}{n!}.$$

Hence the Maclaurin series of e^x is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

which is the series expression for e^x we mentioned previously. We saw back a few lectures ago that this series has infinite radius and interval of convergence (using the ratio test), so we conclude that

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \text{ for all values of } x.$$

In particular, setting $x = 1$ gives

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots,$$

which is an interesting way of expressing the number e. (In fact, as we'll mention soon enough when we talk about Taylor series approximations, this series is *precisely* how computers and calculators come up with decimal expressions for e.)

Now, the Taylor series for e^x centered at -3 can be computed similarly, only that now we evaluate the derivatives we had above at -3. The coefficient of $(x+3)^n$ (which is $(x - \text{center})^n$) in the Taylor series of e^x centered at -3 is then

$$\frac{f^{(n)}(-3)}{n!} = \frac{e^{-3}}{n!},$$

so the Taylor series centered at -3 is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(-3)}{n!}(x+3)^n = \sum_{n=0}^{\infty} \frac{e^{-3}}{n!}(x+3)^n.$$

This series also has infinite radius of convergence, so

$$e^x = \sum_{n=0}^{\infty} \frac{e^{-3}}{n!}(x+3)^n (\text{which holds for all } x)$$

gives another way of representing e^x as a series, only this time as a power series centered at -3 instead of 0. For instance, setting $x = 1$,

$$e = \sum_{n=0}^{\infty} \frac{e^{-3}}{n!} 4^n$$

gives another way of expressing the number e as an infinite sum.

To note one last thing about this example, here we computed the Taylor series of e^x centered at -3 directly using the definition of a Taylor series, but in this case we can also computed it using the Maclaurin series of e^x. Since e^x can be written as

$$e^x = e^{-3} e^{x+3},$$

we can find a series representation for e^x by taking one for e^{x+3} and then multiplying by e^{-3}. Since

$$e^y = \sum_{n=0}^{\infty} \frac{y^n}{n!}$$

based on the Maclaurin series we found above for e^x, setting $y = x + 3$ gives

$$e^{x+3} = \sum_{n=0}^{\infty} \frac{(x+3)^n}{n!},$$

so

$$e^x = e^{-3} e^{x+3} = e^{-3} \sum_{n=0}^{\infty} \frac{1}{n!}(x+3)^n = \sum_{n=0}^{\infty} \frac{e^{-3}}{n!}(x+3)^n,$$

which is precisely the Taylor series for e^x centered at 3 we found before. This is meant to be yet another example of the use of manipulation to turn one series into another.

Example 2. Consider the indefinite integral

$$\int e^{x^2}\, dx.$$

We've mentioned many times in this course that this is not an integral which can be computed directly, so up until now we've haven't been able to say anything more about what this integral actually is. However, the point is that now we *do* have a way of making sense of this integral, since we can represent it as a series! Since

$$e^y = \sum_{n=0}^{\infty} \frac{y^n}{n!},$$

setting $y = x^2$ gives

$$e^{x^2} = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!}.$$

Integrating both sides gives

$$\int e^{x^2}\,dx = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{n!(2n+1)},$$

which is our desired series representation.

Integrals like the one here and others which cannot be computed directly using the integration techniques we spend the first portion of the course studying show up all the time in applications, in fact these types of integrals probably show up more often than integrals which *can* be computed directly. The point is that in these applications the best you can do with such integrals is find a way to represent them as power series, which, as we'll talk about later when discussing approximations, is for most purposes good enough.

Example 3. Finally, we compute the Maclaurin series of $\cos x$. We first need derivatives:

$$f(x) = \cos x,\ f'(x) = -\sin x,\ f''(x) = -\cos x,\ f^{(3)}(x) = \sin x,$$

and after this we repeat these same derivatives: the fourth derivative is $\cos x$, the fifth is $-\sin x$, and so on. Evaluating these are 0 gives

$$f^{(0)}(0) = 1,\ f^{(1)}(0) = 0,\ f^{(2)}(0) = -1,\ f^{(3)}(0) = 0,$$

after which we repeat these values. The Maclauring series of $\cos x$ thus looks like:

$$1 + 0x - \frac{1}{2!}x^2 + \frac{0}{3!}x^3 + \frac{1}{4!}x^4 + \cdots = 1 - \frac{x^2}{2!} + \frac{1}{x^4}4! - \frac{1}{6!}x^6 + \cdots.$$

Notice thus the coefficient of any odd power of x in the Maclaurin series is 0, since all of these coefficients come from evaluating $\pm\sin x$ at $x = 0$, and the coefficients of the even powers are all ± 1 divided by an even factorial, since all of these come from evaluating $\pm\cos x$ at $x = 0$.

When writing the Maclaurin series in a nice way it thus makes sense to write it to include only the terms involving an even power of x, since the other terms will all be zero anyway. We thus write the Maclaurin series of $\cos x$ as

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Compare this with the infinite sum

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

we wrote out earlier: the point is that the $(-1)^n$ portion of our Maclaurin series expression describes the alternating signs, and the $\frac{x^{2n}}{(2n)!}$ portion describes the fact that we only have even powers of x, and such an even power of x is divided by the factorial of the even number describing that power itself. This Maclaurin series of $\cos x$ is one you should have engrained in your minds and know by heart.

Lecture 24: More on Taylor Series

Warm-Up 1. We compute the Maclaurin series of $f(x) = \sin x$. We have:

$$f(x) = \sin x,\ f'(x) = \cos x,\ f''(x) = -\sin x,\ f^{(3)}(x) = -\cos x,$$

and after this the derivatives start repeating: $f^{(4)}(x) = \sin x,\ f^{(5)}(x) = \cos x$, and so on. Evaluating at 0 gives:

$$f(0) = 0,\ f'(0) = 1,\ f''(0) = 0,\ f^{(3)}(0) = -1, \text{ and so on.}$$

The Maclaurin series thus looks like:

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n &= 0 + x + 0 - \frac{x^3}{3!} + 0 + \frac{x^5}{5!} + \cdots \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots .\end{aligned}$$

The point is that $f^{(n)}(0) = 0$ when n is even, so there are no terms in the Maclaurin series involving an even power of x, and the coefficients of the odd-power of x terms alternate between 1 and -1. To write out a more succinct expression for the Maclaurin series we use $(-1)^n$ to deal with the alternating coefficients and use x^{2n+1} to denote odd powers of x:

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

Now, we claim this series as infinite radius of convergence. We have:

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{x^{2(n+1)+1}}{[2(n+1)+1]!} \frac{(2n+1)!}{x^{2n+1}} = \lim_{n\to\infty} \frac{x^2}{(2n+3)(2n+2)} = 0$$

for any value of x. Thus the given series converges for all values of x by the ratio test, and we have

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

with interval of convergence $(-\infty, \infty)$.

Warm-Up 2. With the Maclaurin series for $\sin x$ we can now compute series representations of integrals such as

$$\int x^3 \sin(x^2)\, dx.$$

Since

$$\sin y = \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)!},$$

setting $y = x^2$ gives

$$\sin(x^2) = \sum_{n=0}^{\infty} (-1)^n \frac{(x^2)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+2}}{(2n+1)!}.$$

Multiplying through by x^3 gives

$$x^3 \sin(x^2) = \sum_{n=0}^{\infty} (-1)^n x^3 \frac{x^{4n+2}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+5}}{(2n+1)!},$$

and integrating gives

$$\int x^3 \sin(x^2)\, dx = \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+6}}{(2n+1)!(4n+6)} + C$$

as the desired series representation.

Warm-Up 3. We determine the Taylor series of $\frac{1}{x}$ centered at -3. We first compute derivatives:

$$f(x) = \frac{1}{x}, f'(x) = -\frac{1}{x^2},\ f''(x) = \frac{2}{x^3},\ f^{(3)}(x) = -\frac{2 \cdot 3}{x^4},\ f^{(4)}(x) = \frac{2 \cdot 3 \cdot 4}{x^5},$$

and in general

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

Thus

$$f^{(n)}(-3) = (-1)^n \frac{n!}{(-3)^{n+1}} = -\frac{n!}{3^{n+1}},$$

so the Taylor series of $\frac{1}{x}$ centered at -3 is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(-3)}{n!}[x-(-3)]^n = \sum_{n=0}^{\infty} -\frac{n!}{3^{n+1}n!}(x+3)^n = \sum_{n=0}^{\infty} -\frac{1}{3^{n+1}}(x+3)^n.$$

For good measure, let us determine the radius of convergence of this Taylor series. We have:

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{|x+3|^{n+1}}{3^{n+2}} \frac{3^{n+1}}{|x+3|^n} = \lim_{n\to\infty} \frac{|x+3|}{3} = \frac{|x+3|}{3}.$$

Thus according to the ratio test this series converges when

$$\frac{|x+3|}{3} < 1, \text{ or } |x+3| < 3.$$

Thus the Taylor series for $\frac{1}{x}$ centered at -3 has radius of convergence 3.

Why Taylor series? We now come to the question: why we should care about Taylor series at all? In particular, what's with the strange looking coefficients

$$\frac{f^{(n)}(a)}{n!}?$$

The entire motivation for this comes from wanting to represent a function $f(x)$ as a power series centered at a:

$$f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots.$$

Setting $x = a$ in this expression gives

$$f(a) = c_0 + c_1 0 + c_2 0 + c_3 0 + \cdots = c_0,$$

so we first conclude that the unknown constant term c_0 *must* be $f(a)$. Now, taking the derivative of our series expression gives:

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \cdots .$$

Setting $x = a$ gives

$$f'(a) = c_1 + 2c_2 0 + 3c_3 0 + \cdots = c_1,$$

so the coefficient c_1 of $x - a$ in our series expression must be $f'(a)$. Taking another derivative gives

$$f''(x) = 2c_2 + 2 \cdot 3c_3(x-a) + \cdots$$

and setting $x = a$ gives

$$f''(a) = 2c_2, \text{ so } c_2 = \frac{f''(a)}{2}.$$

Taking another derivative gives

$$f^{(3)}(x) = 2 \cdot 3c_2 + \text{stuff involving } x - a$$

and setting $x = a$ gives

$$f^{(3)}(a) = 2 \cdot 3c_2, \text{ so } c_2 = \frac{f^{(3)}(a)}{3!}.$$

In general, the n-th derivative of

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots$$

looks like

$$f^{(n)}(x) = n!c_n + \text{stuff involving } x - a,$$

and setting $x = a$ gives

$$f^{(n)}(a) = n!c_n + 0, \text{ so } c_n = \frac{f^{(n)}(a)}{n!}.$$

The point of all this is to say that if we want to express f as a power series centered at a:

$$f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n,$$

it turns out that the coefficients *must* be given by the coefficients in the Taylor series of f centered at a. Thus, we have no choice: Taylor series give us the *only* way of expressing a function as a power series if our aim is to indeed express a function as a power series. BOOM!

How to compute $\sin 1$. Recall the Maclaurin series for $\sin x$ we derived in the first Warm-Up:

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!},$$

which as infinite radius of convergence. Setting $x = 1$ gives the equality

$$\sin 1 = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} = 1 - \frac{1}{3!} + \frac{1}{5!} - \frac{1}{7!} + \cdots.$$

Of course, this infinite sum is not possible to compute directly since we can't sit down and literally add up infinitely many quantities. However, taking finite portions of this sum *will* give approximations to $\sin 1$:

$$\sin 1 \approx 1 - \frac{1}{3!} + \frac{1}{5!}$$

for instance. And this is the point: approximating values of functions using power series is *the* way modern computations of numbers actually come about.

For instance, when I plug $\sin 1$ into my calculator I get:

$$\sin 1 \approx 0.8414709848.$$

How did my calculator come up with this value? The answer is that whomever wrote my phone's calculator software programmed the Maclaurin series for $\sin x$ into the phone, and or at least it programmed as many terms of this series as needed to give approximations good enough to include all the decimal places my phone's screen can actually show! In general, the value of $\sin x$ can be approximated incredibly well by taking enough terms in the Maclaurin series of $\sin x$:

$$\sin x \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!}.$$

The expression on the right is the type of thing my calculator has programmed into its memory, and plugging in various values of x gives very good approximations to $\sin x$.

Taylor polynomials. In general, when we have a function as expressed as power series centered at a:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

we can approximate f using the first however-many terms we want of this series, and the idea is that taking more and more terms gives better and better approximations. The expression obtained by taking the terms up to the n-th one (i.e. the term involving the n-th power of $x - a$) is called the *n-th degree Taylor polynomial of f centered at a*:

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

These polynomials form the basis for all modern numerical computations in TONS of applications.

For instance, recall that the Maclaurin series of e^x is:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

The first degree Taylor polynomial centered at 0 gives the approximation

$$e^x \approx 1 + x,$$

the second degree Taylor polynomial gives a better approximation

$$e^x \approx 1 + x + \frac{x^2}{2!},$$

the third degree gives the even better approximation

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!},$$

and so on. Again, next time we'll talk about how to determine just how good these approximations actually are. For now, let me point out that the Wikipedia page for "Taylor series" has a nice animation showing graphically that these polynomials indeed give better and better approximations: the graph for e^x is drawn first, then the graph of $1+x$, then the graph of $1+x+\frac{x^2}{2!}$, and so on, and at each step you can actually see that the Taylor polynomial graphs are getting closer and closer to the actual graph of e^x. Check it out!

Lecture 25: Taylor Polynomials and Approximations

Warm-Up. We compute the 4-th degree Taylor polynomial of $\frac{1}{3-x}$ centered at 1. We have:

$$f(x) = \frac{1}{3-x},\ f'(x) = \frac{1}{(3-x)^2},\ f''(x) = \frac{2}{(3-x)^3},\ f^{(3)}(x) = \frac{2\cdot 3}{(3-x)^4},\ f^{(4)}(x) = \frac{2\cdot 3\cdot 4}{(3-x)^5}.$$

Thus the required Taylor polynomial is

$$f(1)+f'(1)(x-a)+\frac{f''(1)}{2}(x-1)^2+\frac{f^{(3)}(1)}{3!}(x-1)^3+\frac{f^{(4)}(1)}{4!}((x-1)^4 = \frac{1}{2}+\frac{1}{4}(x-1)+\frac{1}{8}(x-1)^2+\frac{1}{16}(x-1)^3+\frac{1}{32}(x-1)^4.$$

Approximations. Recall the idea that the Taylor polynomials of a function provide approximations to that function:

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Let us denote this Taylor polynomial by $T_n(x)$. The *error* arising when using this polynomial to approximate f is given by the difference

$$f(x) - T_n(x),$$

since this tells us precisely how far off the value $T_n(x)$ is from the value $f(x)$. (Actually, since this difference can be positive or negative, we usually care more about the absolute value of this difference $|f(x)-T_n(x)|$.) The difference $f(x)-T_n(x)$ is also called the n-th order Taylor *remainder*.

The key fact which makes Taylor series incredibly worthwhile is that we can actually express this remainder in terms of the function $f(x)$ itself; namely, it is true that we can write this remainder as

$$f(x) - T_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

for some value of c between x and a. The point is that the error obtained when approximating a function using a Taylor polynomial can be written using the *next higher-order derivative* of f and the next higher-order power of $x-a$. So, if we use the 3-rd degree Taylor polynomial, the error will involve terms of degree (or order) 4, if we use the 6-th degree Taylor polynomial the error will use the 7-order term, and so on. This is the key observation which allows us to get a handle on how good of an approximation we have.

Example 1. The 3-rd, 4-th, and 5-th degree Taylor polynomials (centered at 0) of e^x are:

$$e^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{3!}$$
$$e^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!}$$
$$e^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!}.$$

Setting $x = 1$ gives:

$$e \approx 1 + 1 + \frac{1}{2} + \frac{1}{3!} \approx 2.67$$
$$e \approx 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} \approx 2.708$$
$$e \approx 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} \approx 2.7167.$$

The actual value of e is approximately $e \approx 2.71828$, so we indeed see that the approximations given the Taylor polynomials get better and better as the degree of the polynomial increases.

The 5-th degree approximation above has an error term which can be written as

$$|e - (\text{5-th degree approximation})| = \left| \frac{f^{(6)}(c)}{6!} 1^6 \right| = \frac{e^c}{6!}$$

for some value of c between 0 (the center) and 1 (the value we are plugging in).

Example 2. Consider the integral

$$\int_0^1 \sin(x^2)\, dx,$$

which cannot be computed explicitly. However, since

$$\sin y = \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)!},$$

we have

$$\sin(x^2) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+2}}{(2n+1)!}.$$

Thus

$$\int_0^1 \sin(x^2)\, dx = \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+3}}{(2n+1)!(4n+3)} \Big|_0^1 = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!(4n+3)}.$$

The second degree approximation (going up to $n = 2$) is

$$\int_0^1 \sin(x^2)\, dx \approx \frac{1}{3} - \frac{1}{42} + \frac{1}{1320} \approx 0.31028.$$

Now, how good is this approximation? This requires estimating the error

$$|(\text{actual integral value}) - (\text{second degree approximation})|.$$

We could do this using the Taylor remainder, but since the series we used to derive this approximation is actually an alternating series, we can also use the alternating series estimate we spoke about briefly when discussin alternating series. Recall that for an alternating series $\sum(-1)^n b_n$, the error obtained in approximating the actual value using the n-th partial sum is at most b_{n+1}:

$$|(\text{actual value}) - (b_0 - b_1 + b_2 + \cdots + (-1)^n b_n)| \leq b_{n+1}.$$

In our case, since we used the $n = 2$ partial sum, we get that

$$|(\text{actual integral value}) - (\text{second degree approximation})| \leq \frac{1}{75600},$$

where

$$b_n = \frac{1}{(2n+1)!(4n+3)}, \text{ so } b_3 = \frac{1}{75600}.$$

Since

$$\frac{1}{75600} < \frac{1}{10^5},$$

this implies that the actual value of the integral

$$\int_0^1 \sin(x^2)\, dx$$

and the approximation

$$0.31028$$

using the 2-nd order approximation should agree to at least 4 decimal places. Indeed, the actual value of this integral (obtained from my calculator, which itself uses an even better approximation than the one we're using here) is about

$$0.310268,$$

and low-and-behold this two values do agree to 4 decimal places!

Example 3. Finally, consider the following 4-th degree Taylor approximation of $\cos x$:

$$\cos x \approx 1 - \frac{x^2}{2!} + \frac{x^4}{4!},$$

where the expression on the right is the 4-th degree Taylor polynomial centered at 0. Suppose we use this polynomial to approximate $\cos x$ for values of x in the interval $(-0.5, 0.5)$, or equivalently values of x satisfying $|x| \leq 0.5$. We determine how good of an approximation this is.

Since we approximate using the 4-th degree Taylor polynomial, the error should be given by a term using the 5-th derivative of $f(x) = \cos x$:

$$|\cos x - (\text{4-th degree Taylor polynomial})| = \left|\frac{f^{(5)}(c)}{5!}x^5\right|$$

for some c between 0 and x. Since $f(x) = \cos x$, $f^{(5)}(x) = -\sin c$, so

$$\left|\frac{f^{(5)}(c)}{5!}x^5\right| = \frac{|-\sin c|}{120}|x|^5.$$

Since $|-\sin c| \leq 1$ and we are considering x satisfying $|0.5|^5$, we get that the error is at most:

$$\text{error} \leq \frac{1}{120}(0.5)^5 \approx 0.00026 \leq \frac{1}{10^3} = 0.001.$$

Thus for $|x| \leq 0.5$, $\cos x$ and

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!}$$

are within 0.001 of one another, so their values agree to at least 2 decimal places. As a check: plugging $x = 0.25$ into this Taylor approximation gives a value of

$$0.9689176,$$

while the *actual* value of $\cos 0.25$ is about

$$0.9688912,$$

so indeed these two values agree to at least 2 (in fact more) decimal places. Huzzah, math works!

Summary. And so the course comes to an end. We've covered a lot this quarter, but just to highlight the overarching concept tying everything together: this course was all about infinite summations. Indeed, we've stated at various points that integrals should be viewed as a type of infinite summations, and of course series are also types of infinite summations as well. In more technical terms, an integral is a type of *continuous* summation, while a series is a type of *discrete* summation. The key point is that summations and approximations go hand-in-hand, and really do underlie many computational techniques in various applications, some of which you may no doubt see in future courses or in your careers. I hope you've enjoyed the journey.