CALCULUS II

Paul Dawkins

Table of Contents

Preface

Here are my online notes for my Calculus II course that I teach here at Lamar University. Despite the fact that these are my "class notes", they should be accessible to anyone wanting to learn Calculus II or needing a refresher in some of the topics from the class.

These notes do assume that the reader has a good working knowledge of Calculus I topics including limits, derivatives and basic integration and integration by substitution.

Calculus II tends to be a very difficult course for many students. There are many reasons for this.

The first reason is that this course does require that you have a very good working knowledge of Calculus I. The Calculus I portion of many of the problems tends to be skipped and left to the student to verify or fill in the details. If you don't have good Calculus I skills, and you are constantly getting stuck on the Calculus I portion of the problem, you will find this course very difficult to complete.

The second, and probably larger, reason many students have difficulty with Calculus II is that you will be asked to truly think in this class. That is not meant to insult anyone; it is simply an acknowledgment that you can't just memorize a bunch of formulas and expect to pass the course as you can do in many math classes. There are formulas in this class that you will need to know, but they tend to be fairly general. You will need to understand them, how they work, and more importantly whether they can be used or not. As an example, the first topic we will look at is Integration by Parts. The integration by parts formula is very easy to remember. However, just because you've got it memorized doesn't mean that you can use it. You'll need to be able to look at an integral and realize that integration by parts can be used (which isn't always obvious) and then decide which portions of the integral correspond to the parts in the formula (again, not always obvious).

Finally, many of the problems in this course will have multiple solution techniques and so you'll need to be able to identify all the possible techniques and then decide which will be the easiest technique to use.

So, with all that out of the way let me also get a couple of warnings out of the way to my students who may be here to get a copy of what happened on a day that you missed.

1. Because I wanted to make this a fairly complete set of notes for anyone wanting to learn calculus I have included some material that I do not usually have time to cover in class and because this changes from semester to semester it is not noted here. You will need to find one of your fellow class mates to see if there is something in these notes that wasn't covered in class.

2. In general I try to work problems in class that are different from my notes. However, with Calculus II many of the problems are difficult to make up on the spur of the moment and so in this class my class work will follow these notes fairly close as far as worked problems go. With that being said I will, on occasion, work problems off the top of my head when I can to provide more examples than just those in my notes. Also, I often

don't have time in class to work all of the problems in the notes and so you will find that some sections contain problems that weren't worked in class due to time restrictions.

3. Sometimes questions in class will lead down paths that are not covered here. I try to anticipate as many of the questions as possible in writing these up, but the reality is that I can't anticipate all the questions. Sometimes a very good question gets asked in class that leads to insights that I've not included here. You should always talk to someone who was in class on the day you missed and compare these notes to their notes and see what the differences are.

4. This is somewhat related to the previous three items, but is important enough to merit its own item. THESE NOTES ARE NOT A SUBSTITUTE FOR ATTENDING CLASS!! Using these notes as a substitute for class is liable to get you in trouble. As already noted not everything in these notes is covered in class and often material or insights not in these notes is covered in class.

Outline

Here is a listing and brief description of the material in this set of notes.

Integration Techniques

Introduction

In this chapter we are going to be looking at various integration techniques. There are a fair number of them and some will be easier than others. The point of the chapter is to teach you these new techniques and so this chapter assumes that you've got a fairly good working knowledge of basic integration as well as substitutions with integrals. In fact, most integrals involving "simple" substitutions will not have any of the substitution work shown. It is going to be assumed that you can verify the substitution portion of the integration yourself.

Also, most of the integrals done in this chapter will be indefinite integrals. It is also assumed that once you can do the indefinite integrals you can also do the definite integrals and so to conserve space we concentrate mostly on indefinite integrals. There is one exception to this and that is the Trig Substitution section and in this case there are some subtleties involved with definite integrals that we're going to have to watch out for. Outside of that however, most sections will have at most one definite integral example and some sections will not have any definite integral examples.

Here is a list of topics that are covered in this chapter.

Integration by Parts – Of all the integration techniques covered in this chapter this is probably the one that students are most likely to run into down the road in other classes.

Integrals Involving Trig Functions – In this section we look at integrating certain products and quotients of trig functions.

Trig Substitutions – Here we will look using substitutions involving trig functions and how they can be used to simplify certain integrals.

Partial Fractions – We will use partial fractions to allow us to do integrals involving some rational functions.

Integrals Involving Roots – We will take a look at a substitution that can, on occasion, be used with integrals involving roots.

Integrals Involving Quadratics – In this section we are going to look at some integrals that involve quadratics.

Integration Strategy – We give a general set of guidelines for determining how to evaluate an integral.

Improper Integrals – We will look at integrals with infinite intervals of integration and integrals with discontinuous integrands in this section.

Comparison Test for Improper Integrals – Here we will use the Comparison Test to determine if improper integrals converge or diverge.

Approximating Definite Integrals – There are many ways to approximate the value of a definite integral. We will look at three of them in this section.

Integration by Parts

Let's start off with this section with a couple of integrals that we should already be able to do to get us started. First let's take a look at the following.

$$\int \mathbf{e}^{x}\,dx = \mathbf{e}^{x} + c$$

So, that was simple enough. Now, let's take a look at,

$$\int x\mathbf{e}^{x^2}\,dx$$

To do this integral we'll use the following substitution.

$$u = x^2 \qquad du = 2x\,dx \qquad \Rightarrow \qquad x\,dx = \frac{1}{2}du$$

$$\int x\mathbf{e}^{x^2}\,dx = \frac{1}{2}\int \mathbf{e}^{u}\,du = \frac{1}{2}\mathbf{e}^{u} + c = \frac{1}{2}\mathbf{e}^{x^2} + c$$

Again, simple enough to do provided you remember how to do substitutions. By the way make sure that you can do these kinds of substitutions quickly and easily. From this point on we are going to be doing these kinds of substitutions in our head. If you have to stop and write these out with every problem you will find that it will take you significantly longer to do these problems.

Now, let's look at the integral that we really want to do.

$$\int x\mathbf{e}^{6x}\,dx$$

If we just had an x by itself or $\mathbf{e}^{6x}$ by itself we could do the integral easily enough. But, we don't have them by themselves, they are instead multiplied together.

There is no substitution that we can use on this integral that will allow us to do the integral. So, at this point we don't have the knowledge to do this integral.

To do this integral we will need to use integration by parts so let's derive the integration by parts formula. We'll start with the product rule.

$$\left(f\,g\right)' = f'\,g + f\,g'$$

Now, integrate both sides of this.

$$\int \left(f\,g\right)'\,dx = \int f'\,g + f\,g'\,dx$$

The left side is easy enough to integrate and we'll split up the right side of the integral.

$$fg = \int f'\,g\,dx + \int f\,g'\,dx$$

Note that technically we should have had a constant of integration show up on the left side after doing the integration. We can drop it at this point since other constants of integration will be showing up down the road and they would just end up absorbing this one.

Finally, rewrite the formula as follows and we arrive at the integration by parts formula.

$$\int f\, g'\, dx = fg - \int f'\, g\, dx$$

This is not the easiest formula to use however. So, let's do a couple of substitutions.

$$u = f(x) \qquad v = g(x)$$
$$du = f'(x)\,dx \qquad dv = g'(x)\,dx$$

Both of these are just the standard Calc I substitutions that hopefully you are used to by now. Don't get excited by the fact that we are using two substitutions here. They will work the same way.

Using these substitutions gives us the formula that most people think of as the integration by parts formula.

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

To use this formula we will need to identify *u* and *dv*, compute *du* and *v* and then use the formula. Note as well that computing *v* is very easy. All we need to do is integrate *dv*.

$$v = \int dv$$

So, let's take a look at the integral above that we mentioned we wanted to do.

Example 1 Evaluate the following integral.

$$\int x\mathbf{e}^{6x}\, dx$$

Solution

So, on some level, the problem here is the *x* that is in front of the exponential. If that wasn't there we could do the integral. Notice as well that in doing integration by parts anything that we choose for *u* will be differentiated. So, it seems that choosing $u = x$ will be a good choice since upon differentiating the *x* will drop out.

Now that we've chosen *u* we know that *dv* will be everything else that remains. So, here are the choices for *u* and *dv* as well as *du* and *v*.

$$u = x \qquad dv = \mathbf{e}^{6x}\, dx$$
$$du = dx \qquad v = \int \mathbf{e}^{6x}\, dx = \frac{1}{6}\mathbf{e}^{6x}$$

The integral is then,

$$\begin{aligned}\int x\mathbf{e}^{6x}\, dx &= \frac{x}{6}\mathbf{e}^{6x} - \int \frac{1}{6}\mathbf{e}^{6x}\, dx \\ &= \frac{x}{6}\mathbf{e}^{6x} - \frac{1}{36}\mathbf{e}^{6x} + c\end{aligned}$$

Once we have done the last integral in the problem we will add in the constant of integration to get our final answer.

Next, let's take a look at integration by parts for definite integrals. The integration by parts formula for definite integrals is,

Integration by Parts, Definite Integrals

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

Note that the $uv\Big|_a^b$ in the first term is just the standard integral evaluation notation that you should be familiar with at this point. All we do is evaluate the term, uv in this case, at b then subtract off the evaluation of the term at a.

At some level we don't really need a formula here because we know that when doing definite integrals all we need to do is do the indefinite integral and then do the evaluation.

Let's take a quick look at a definite integral using integration by parts.

Example 2 Evaluate the following integral.

$$\int_{-1}^{2} x\mathbf{e}^{6x}\,dx$$

Solution

This is the same integral that we looked at in the first example so we'll use the same u and dv to get,

$$\begin{aligned}\int_{-1}^{2} x\mathbf{e}^{6x}\,dx &= \frac{x}{6}\mathbf{e}^{6x}\Big|_{-1}^{2} - \frac{1}{6}\int_{-1}^{2}\mathbf{e}^{6x}\,dx\\ &= \frac{x}{6}\mathbf{e}^{6x}\Big|_{-1}^{2} - \frac{1}{36}\mathbf{e}^{6x}\Big|_{-1}^{2}\\ &= \frac{11}{36}\mathbf{e}^{12} + \frac{7}{36}\mathbf{e}^{-6}\end{aligned}$$

Since we need to be able to do the indefinite integral in order to do the definite integral and doing the definite integral amounts to nothing more than evaluating the indefinite integral at a couple of points we will concentrate on doing indefinite integrals in the rest of this section. In fact, throughout most of this chapter this will be the case. We will be doing far more indefinite integrals than definite integrals.

Let's take a look at some more examples.

Example 3 Evaluate the following integral.

$$\int (3t+5)\cos\left(\frac{t}{4}\right)dt$$

Solution

There are two ways to proceed with this example. For many, the first thing that they try is multiplying the cosine through the parenthesis, splitting up the integral and then doing integration by parts on the first integral.

While that is a perfectly acceptable way of doing the problem it's more work than we really need to do. Instead of splitting the integral up let's instead use the following choices for u and dv.

$$u = 3t + 5 \qquad dv = \cos\left(\frac{t}{4}\right)dt$$

$$du = 3\,dt \qquad v = 4\sin\left(\frac{t}{4}\right)$$

The integral is then,

$$\begin{aligned}\int (3t+5)\cos\left(\frac{t}{4}\right)dt &= 4(3t+5)\sin\left(\frac{t}{4}\right) - 12\int \sin\left(\frac{t}{4}\right)dt \\ &= 4(3t+5)\sin\left(\frac{t}{4}\right) + 48\cos\left(\frac{t}{4}\right) + c\end{aligned}$$

Notice that we pulled any constants out of the integral when we used the integration by parts formula. We will usually do this in order to simplify the integral a little.

Example 4 Evaluate the following integral.

$$\int w^2 \sin(10w)\,dw$$

Solution

For this example we'll use the following choices for u and dv.

$$u = w^2 \qquad dv = \sin(10w)\,dw$$

$$du = 2w\,dw \qquad v = -\frac{1}{10}\cos(10w)$$

The integral is then,

$$\int w^2 \sin(10w)\,dw = -\frac{w^2}{10}\cos(10w) + \frac{1}{5}\int w\cos(10w)\,dw$$

In this example, unlike the previous examples, the new integral will also require integration by parts. For this second integral we will use the following choices.

$$u = w \qquad dv = \cos(10w)\,dw$$

$$du = dw \qquad v = \frac{1}{10}\sin(10w)$$

So, the integral becomes,

$$\begin{aligned}\int w^2 \sin(10w)\,dw &= -\frac{w^2}{10}\cos(10w) + \frac{1}{5}\left(\frac{w}{10}\sin(10w) - \frac{1}{10}\int \sin(10w)\,dw\right) \\ &= -\frac{w^2}{10}\cos(10w) + \frac{1}{5}\left(\frac{w}{10}\sin(10w) + \frac{1}{100}\cos(10w)\right) + c \\ &= -\frac{w^2}{10}\cos(10w) + \frac{w}{50}\sin(10w) + \frac{1}{500}\cos(10w) + c\end{aligned}$$

Be careful with the coefficient on the integral for the second application of integration by parts. Since the integral is multiplied by $\frac{1}{5}$ we need to make sure that the results of actually doing the integral are also multiplied by $\frac{1}{5}$. Forgetting to do this is one of the more common mistakes with integration by parts problems.

As this last example has shown us, we will sometimes need more than one application of integration by parts to completely evaluate an integral. This is something that will happen so don't get excited about it when it does.

In this next example we need to acknowledge an important point about integration techniques. Some integrals can be done in using several different techniques. That is the case with the integral in the next example.

Example 5 Evaluate the following integral

$$\int x\sqrt{x+1}\,dx$$

(a) Using Integration by Parts. [Solution]

(b) Using a standard Calculus I substitution. [Solution]

Solution

(a) Evaluate using Integration by Parts.

First notice that there are no trig functions or exponentials in this integral. While a good many integration by parts integrals will involve trig functions and/or exponentials not all of them will so don't get too locked into the idea of expecting them to show up.

In this case we'll use the following choices for u and dv.

$$u = x \qquad dv = \sqrt{x+1}\,dx$$

$$du = dx \qquad v = \frac{2}{3}\left(x+1\right)^{\frac{3}{2}}$$

The integral is then,

$$\begin{aligned}\int x\sqrt{x+1}\,dx &= \frac{2}{3}x\left(x+1\right)^{\frac{3}{2}} - \frac{2}{3}\int\left(x+1\right)^{\frac{3}{2}}dx \\ &= \frac{2}{3}x\left(x+1\right)^{\frac{3}{2}} - \frac{4}{15}\left(x+1\right)^{\frac{5}{2}} + c\end{aligned}$$

[Return to Problems]

(b) Evaluate Using a standard Calculus I substitution.

Now let's do the integral with a substitution. We can use the following substitution.

$$u = x+1 \qquad x = u-1 \qquad du = dx$$

Notice that we'll actually use the substitution twice, once for the quantity under the square root and once for the x in front of the square root. The integral is then,

$$\begin{aligned}\int x\sqrt{x+1}\,dx &= \int\left(u-1\right)\sqrt{u}\,du \\ &= \int u^{\frac{3}{2}} - u^{\frac{1}{2}}\,du \\ &= \frac{2}{5}u^{\frac{5}{2}} - \frac{2}{3}u^{\frac{3}{2}} + c \\ &= \frac{2}{5}\left(x+1\right)^{\frac{5}{2}} - \frac{2}{3}\left(x+1\right)^{\frac{3}{2}} + c\end{aligned}$$

[Return to Problems]

So, we used two different integration techniques in this example and we got two different answers. The obvious question then should be : Did we do something wrong?

Actually, we didn't do anything wrong. We need to remember the following fact from Calculus I.

$$\text{If} \quad f'(x) = g'(x) \quad \text{then} \quad f(x) = g(x) + c$$

In other words, if two functions have the same derivative then they will differ by no more than a constant. So, how does this apply to the above problem? First define the following,

$$f'(x) = g'(x) = x\sqrt{x+1}$$

Then we can compute $f(x)$ and $g(x)$ by integrating as follows,

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

We'll use integration by parts for the first integral and the substitution for the second integral. Then according to the fact $f(x)$ and $g(x)$ should differ by no more than a constant. Let's verify this and see if this is the case. We can verify that they differ by no more than a constant if we take a look at the difference of the two and do a little algebraic manipulation and simplification.

$$\begin{aligned}\left(\frac{2}{3}x(x+1)^{\frac{3}{2}} - \frac{4}{15}(x+1)^{\frac{5}{2}}\right) - \left(\frac{2}{5}(x+1)^{\frac{5}{2}} - \frac{2}{3}(x+1)^{\frac{3}{2}}\right) \\ &= (x+1)^{\frac{3}{2}}\left(\frac{2}{3}x - \frac{4}{15}(x+1) - \frac{2}{5}(x+1) + \frac{2}{3}\right) \\ &= (x+1)^{\frac{3}{2}}(0) \\ &= 0\end{aligned}$$

So, in this case it turns out the two functions are exactly the same function since the difference is zero. Note that this won't always happen. Sometimes the difference will yield a nonzero constant. For an example of this check out the Constant of Integration section in my Calculus I notes.

So just what have we learned? First, there will, on occasion, be more than one method for evaluating an integral. Secondly, we saw that different methods will often lead to different answers. Last, even though the answers are different it can be shown, sometimes with a lot of work, that they differ by no more than a constant.

When we are faced with an integral the first thing that we'll need to decide is if there is more than one way to do the integral. If there is more than one way we'll then need to determine which method we should use. The general rule of thumb that I use in my classes is that you should use the method that *you* find easiest. This may not be the method that others find easiest, but that doesn't make it the wrong method.

One of the more common mistakes with integration by parts is for people to get too locked into perceived patterns. For instance, all of the previous examples used the basic pattern of taking *u* to be the polynomial that sat in front of another function and then letting *dv* be the other function. This will not always happen so we need to be careful and not get locked into any patterns that we think we see.

Let's take a look at some integrals that don't fit into the above pattern.

Example 6 Evaluate the following integral.

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

Solution

So, unlike any of the other integral we've done to this point there is only a single function in the integral and no polynomial sitting in front of the logarithm.

The first choice of many people here is to try and fit this into the pattern from above and make the following choices for *u* and *dv*.

$$u = 1 \qquad\qquad dv = \ln x\,dx$$

This leads to a real problem however since that means *v* must be,

$$v = \int \ln x\,dx$$

In other words, we would need to know the answer ahead of time in order to actually do the problem. So, this choice simply won't work. Also notice that with this choice we'd get that $du = 0$ which also causes problems and is another reason why this choice will not work.

Therefore, if the logarithm doesn't belong in the *dv* it must belong instead in the *u*. So, let's use the following choices instead

$$u = \ln x \qquad\qquad dv = dx$$

$$du = \frac{1}{x}dx \qquad\qquad v = x$$

The integral is then,

$$\begin{aligned}\int \ln x\,dx &= x\ln x - \int \frac{1}{x}x\,dx \\ &= x\ln x - \int dx \\ &= x\ln x - x + c\end{aligned}$$

Example 7 Evaluate the following integral.

$$\int x^5\sqrt{x^3+1}\,dx$$

Solution

So, if we again try to use the pattern from the first few examples for this integral our choices for *u* and *dv* would probably be the following.

$$u = x^5 \qquad\qquad dv = \sqrt{x^3+1}\,dx$$

However, as with the previous example this won't work since we can't easily compute *v*.

$$v = \int \sqrt{x^3+1}\,dx$$

This is not an easy integral to do. However, notice that if we had an x^2 in the integral along with the root we could very easily do the integral with a substitution. Also notice that we do have a lot of x's floating around in the original integral. So instead of putting all the x's (outside of the root) in the u let's split them up as follows.

$$u = x^3 \qquad dv = x^2\sqrt{x^3+1}\,dx$$

$$du = 3x^2 dx \qquad v = \frac{2}{9}\left(x^3+1\right)^{\frac{3}{2}}$$

We can now easily compute v and after using integration by parts we get,

$$\begin{aligned}\int x^5\sqrt{x^3+1}\,dx &= \frac{2}{9}x^3\left(x^3+1\right)^{\frac{3}{2}} - \frac{2}{3}\int x^2\left(x^3+1\right)^{\frac{3}{2}}dx \\ &= \frac{2}{9}x^3\left(x^3+1\right)^{\frac{3}{2}} - \frac{4}{45}\left(x^3+1\right)^{\frac{5}{2}} + c\end{aligned}$$

So, in the previous two examples we saw cases that didn't quite fit into any perceived pattern that we might have gotten from the first couple of examples. This is always something that we need to be on the lookout for with integration by parts.

Let's take a look at another example that also illustrates another integration technique that sometimes arises out of integration by parts problems.

Example 8 Evaluate the following integral.

$$\int \mathbf{e}^{\theta}\cos\theta\,d\theta$$

Solution

Okay, to this point we've always picked u in such a way that upon differentiating it would make that portion go away or at the very least put it the integral into a form that would make it easier to deal with. In this case no matter which part we make u it will never go away in the differentiation process.

It doesn't much matter which we choose to be u so we'll choose in the following way. Note however that we could choose the other way as well and we'll get the same result in the end.

$$u = \cos\theta \qquad dv = \mathbf{e}^{\theta}\,d\theta$$

$$du = -\sin\theta\,d\theta \qquad v = \mathbf{e}^{\theta}$$

The integral is then,

$$\int \mathbf{e}^{\theta}\cos\theta\,d\theta = \mathbf{e}^{\theta}\cos\theta + \int \mathbf{e}^{\theta}\sin\theta\,d\theta$$

So, it looks like we'll do integration by parts again. Here are our choices this time.

$$u = \sin\theta \qquad dv = \mathbf{e}^{\theta}\,d\theta$$

$$du = \cos\theta\,d\theta \qquad v = \mathbf{e}^{\theta}$$

The integral is now,

$$\int \mathbf{e}^{\theta}\cos\theta\,d\theta = \mathbf{e}^{\theta}\cos\theta + \mathbf{e}^{\theta}\sin\theta - \int \mathbf{e}^{\theta}\cos\theta\,d\theta$$

Now, at this point it looks like we're just running in circles. However, notice that we now have the same integral on both sides and on the right side it's got a minus sign in front of it. This means that we can add the integral to both sides to get,

$$2\int \mathbf{e}^{\theta} \cos\theta \, d\theta = \mathbf{e}^{\theta} \cos\theta + \mathbf{e}^{\theta} \sin\theta$$

All we need to do now is divide by 2 and we're done. The integral is,

$$\int \mathbf{e}^{\theta} \cos\theta \, d\theta = \frac{1}{2}\left(\mathbf{e}^{\theta} \cos\theta + \mathbf{e}^{\theta} \sin\theta\right) + c$$

Notice that after dividing by the two we add in the constant of integration at that point.

This idea of integrating until you get the same integral on both sides of the equal sign and then simply solving for the integral is kind of nice to remember. It doesn't show up all that often, but when it does it may be the only way to actually do the integral.

We've got one more example to do. As we will see some problems could require us to do integration by parts numerous times and there is a short hand method that will allow us to do multiple applications of integration by parts quickly and easily.

Example 9 Evaluate the following integral.

$$\int x^4 \mathbf{e}^{\frac{x}{2}} \, dx$$

Solution

We start off by choosing *u* and *dv* as we always would. However, instead of computing *du* and *v* we put these into the following table. We then differentiate down the column corresponding to *u* until we hit zero. In the column corresponding to *dv* we integrate once for each entry in the first column. There is also a third column which we will explain in a bit and it always starts with a "+" and then alternates signs as shown.

x^4	$\mathbf{e}^{\frac{x}{2}}$	+
$4x^3$	$2\mathbf{e}^{\frac{x}{2}}$	−
$12x^2$	$4\mathbf{e}^{\frac{x}{2}}$	+
$24x$	$8\mathbf{e}^{\frac{x}{2}}$	−
24	$16\mathbf{e}^{\frac{x}{2}}$	+
0	$32\mathbf{e}^{\frac{x}{2}}$	−

Now, multiply along the diagonals shown in the table. In front of each product put the sign in the third column that corresponds to the "*u*" term for that product. In this case this would give,

$$\int x^4 \mathbf{e}^{\frac{x}{2}}\,dx = \left(x^4\right)\left(2\mathbf{e}^{\frac{x}{2}}\right) - \left(4x^3\right)\left(4\mathbf{e}^{\frac{x}{2}}\right) + \left(12x^2\right)\left(8\mathbf{e}^{\frac{x}{2}}\right) - \left(24x\right)\left(16\mathbf{e}^{\frac{x}{2}}\right) + \left(24\right)\left(32\mathbf{e}^{\frac{x}{2}}\right)$$

$$= 2x^4\mathbf{e}^{\frac{x}{2}} - 16x^3\mathbf{e}^{\frac{x}{2}} + 96x^2\mathbf{e}^{\frac{x}{2}} - 384x\mathbf{e}^{\frac{x}{2}} + 768\mathbf{e}^{\frac{x}{2}} + c$$

We've got the integral. This is much easier than writing down all the various *u*'s and *dv*'s that we'd have to do otherwise.

So, in this section we've seen how to do integration by parts. In your later math classes this is liable to be one of the more frequent integration techniques that you'll encounter.

It is important to not get too locked into patterns that you may think you've seen. In most cases any pattern that you think you've seen can (and will be) violated at some point in time. Be careful!

Also, don't forget the shorthand method for multiple applications of integration by parts problems. It can save you a fair amount of work on occasion.

Integrals Involving Trig Functions

In this section we are going to look at quite a few integrals involving trig functions and some of the techniques we can use to help us evaluate them. Let's start off with an integral that we should already be able to do.

$$\begin{aligned}\int \cos x \sin^5 x\,dx &= \int u^5\,du \qquad \text{using the substitution } u = \sin x\\ &= \frac{1}{6}\sin^6 x + c\end{aligned}$$

This integral is easy to do with a substitution because the presence of the cosine, however, what about the following integral.

Example 1 Evaluate the following integral.

$$\int \sin^5 x\,dx$$

Solution

This integral no longer has the cosine in it that would allow us to use the substitution that we used above. Therefore, that substitution won't work and we are going to have to find another way of doing this integral.

Let's first notice that we could write the integral as follows,

$$\int \sin^5 x\,dx = \int \sin^4 x \sin x\,dx = \int \left(\sin^2 x\right)^2 \sin x\,dx$$

Now recall the trig identity,

$$\cos^2 x + \sin^2 x = 1 \qquad \Rightarrow \qquad \sin^2 x = 1 - \cos^2 x$$

With this identity the integral can be written as,

$$\int \sin^5 x\,dx = \int \left(1 - \cos^2 x\right)^2 \sin x\,dx$$

and we can now use the substitution $u = \cos x$. Doing this gives us,

$$\begin{aligned}\int \sin^5 x\,dx &= -\int \left(1 - u^2\right)^2 du\\ &= -\int 1 - 2u^2 + u^4\,du\\ &= -\left(u - \frac{2}{3}u^3 + \frac{1}{5}u^5\right) + c\\ &= -\cos x + \frac{2}{3}\cos^3 x - \frac{1}{5}\cos^5 x + c\end{aligned}$$

So, with a little rewriting on the integrand we were able to reduce this to a fairly simple substitution.

Notice that we were able to do the rewrite that we did in the previous example because the exponent on the sine was odd. In these cases all that we need to do is strip out one of the sines.

The exponent on the remaining sines will then be even and we can easily convert the remaining sines to cosines using the identity,

$$\cos^2 x + \sin^2 x = 1 \tag{1}$$

If the exponent on the sines had been even this would have been difficult to do. We could strip out a sine, but the remaining sines would then have an odd exponent and while we could convert them to cosines the resulting integral would often be even more difficult than the original integral in most cases.

Let's take a look at another example.

Example 2 Evaluate the following integral.

$$\int \sin^6 x \cos^3 x\,dx$$

Solution

So, in this case we've got both sines and cosines in the problem and in this case the exponent on the sine is even while the exponent on the cosine is odd. So, we can use a similar technique in this integral. This time we'll strip out a cosine and convert the rest to sines.

$$\begin{aligned}\int \sin^6 x \cos^3 x\,dx &= \int \sin^6 x \cos^2 x \cos x\,dx \\ &= \int \sin^6 x \left(1-\sin^2 x\right)\cos x\,dx \qquad\qquad u = \sin x \\ &= \int u^6\left(1-u^2\right)du \\ &= \int u^6 - u^8\,du \\ &= \frac{1}{7}\sin^7 x - \frac{1}{9}\sin^9 x + c\end{aligned}$$

At this point let's pause for a second to summarize what we've learned so far about integrating powers of sine and cosine.

$$\int \sin^n x \cos^m x\,dx \tag{2}$$

In this integral if the exponent on the sines (n) is odd we can strip out one sine, convert the rest to cosines using (1) and then use the substitution $u = \cos x$. Likewise, if the exponent on the cosines (m) is odd we can strip out one cosine and convert the rest to sines and the use the substitution $u = \sin x$.

Of course, if both exponents are odd then we can use either method. However, in these cases it's usually easier to convert the term with the smaller exponent.

The one case we haven't looked at is what happens if both of the exponents are even? In this case the technique we used in the first couple of examples simply won't work and in fact there really isn't any one set method for doing these integrals. Each integral is different and in some cases there will be more than one way to do the integral.

With that being said most, if not all, of integrals involving products of sines and cosines in which both exponents are even can be done using one or more of the following formulas to rewrite the integrand.

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

The first two formulas are the standard half angle formula from a trig class written in a form that will be more convenient for us to use. The last is the standard double angle formula for sine, again with a small rewrite.

Let's take a look at an example.

Example 3 Evaluate the following integral.

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

Solution

As noted above there are often more than one way to do integrals in which both of the exponents are even. This integral is an example of that. There are at least two solution techniques for this problem. We will do both solutions starting with what is probably the harder of the two, but it's also the one that many people see first.

Solution 1

In this solution we will use the two half angle formulas above and just substitute them into the integral.

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

So, we still have an integral that can't be completely done, however notice that we have managed to reduce the integral down to just one term causing problems (a cosine with an even power) rather than two terms causing problems.

In fact to eliminate the remaining problem term all that we need to do is reuse the first half angle formula given above.

$$\begin{aligned}\int \sin^2 x \cos^2 x \, dx &= \frac{1}{4}\int 1-\frac{1}{2}\left(1+\cos\left(4x\right)\right) dx \\ &= \frac{1}{4}\int \frac{1}{2}-\frac{1}{2}\cos\left(4x\right) dx \\ &= \frac{1}{4}\left(\frac{1}{2}x-\frac{1}{8}\sin\left(4x\right)\right)+c \\ &= \frac{1}{8}x-\frac{1}{32}\sin\left(4x\right)+c\end{aligned}$$

So, this solution required a total of three trig identities to complete.

Solution 2
In this solution we will use the half angle formula to help simplify the integral as follows.

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

Now, we use the double angle formula for sine to reduce to an integral that we can do.

$$\begin{aligned}\int \sin^2 x \cos^2 x \, dx &= \frac{1}{8}\int 1 - \cos(4x)\, dx \\ &= \frac{1}{8}x - \frac{1}{32}\sin(4x) + c\end{aligned}$$

This method required only two trig identities to complete.

Notice that the difference between these two methods is more one of "messiness". The second method is not appreciably easier (other than needing one less trig identity) it is just not as messy and that will often translate into an "easier" process.

In the previous example we saw two different solution methods that gave the same answer. Note that this will not always happen. In fact, more often than not we will get different answers. However, as we discussed in the Integration by Parts section, the two answers will differ by no more than a constant.

In general when we have products of sines and cosines in which both exponents are even we will need to use a series of half angle and/or double angle formulas to reduce the integral into a form that we can integrate. Also, the larger the exponents the more we'll need to use these formulas and hence the messier the problem.

Sometimes in the process of reducing integrals in which both exponents are even we will run across products of sine and cosine in which the arguments are different. These will require one of the following formulas to reduce the products to integrals that we can do.

$$\sin\alpha\cos\beta = \frac{1}{2}\left[\sin(\alpha-\beta) + \sin(\alpha+\beta)\right]$$

$$\sin\alpha\sin\beta = \frac{1}{2}\left[\cos(\alpha-\beta) - \cos(\alpha+\beta)\right]$$

$$\cos\alpha\cos\beta = \frac{1}{2}\left[\cos(\alpha-\beta) + \cos(\alpha+\beta)\right]$$

Let's take a look at an example of one of these kinds of integrals.

Example 4 Evaluate the following integral.

$$\int \cos(15x)\cos(4x)\,dx$$

Solution

This integral requires the last formula listed above.

$$\begin{aligned}\int \cos(15x)\cos(4x)\,dx &= \frac{1}{2}\int \cos(11x)+\cos(19x)\,dx \\ &= \frac{1}{2}\left(\frac{1}{11}\sin(11x)+\frac{1}{19}\sin(19x)\right)+c\end{aligned}$$

Okay, at this point we've covered pretty much all the possible cases involving products of sines and cosines. It's now time to look at integrals that involve products of secants and tangents.

This time, let's do a little analysis of the possibilities before we just jump into examples. The general integral will be,

$$\int \sec^n x\tan^m x\,dx \tag{3}$$

The first thing to notice is that we can easily convert even powers of secants to tangents and even powers of tangents to secants by using a formula similar to (1). In fact, the formula can be derived from (1) so let's do that.

$$\sin^2 x+\cos^2 x=1$$

$$\frac{\sin^2 x}{\cos^2 x}+\frac{\cos^2 x}{\cos^2 x}=\frac{1}{\cos^2 x}$$

$$\tan^2 x+1=\sec^2 x \tag{4}$$

Now, we're going to want to deal with (3) similarly to how we dealt with (2). We'll want to eventually use one of the following substitutions.

$$\begin{aligned} u &= \tan x \qquad & du &= \sec^2 x\,dx \\ u &= \sec x \qquad & du &= \sec x\tan x\,dx\end{aligned}$$

So, if we use the substitution $u=\tan x$ we will need two secants left for the substitution to work. This means that if the exponent on the secant (n) is even we can strip two out and then convert the remaining secants to tangents using (4).

Next, if we want to use the substitution $u=\sec x$ we will need one secant and one tangent left over in order to use the substitution. This means that if the exponent on the tangent (m) is odd and we have at least one secant in the integrand we can strip out one of the tangents along with one of the secants of course. The tangent will then have an even exponent and so we can use (4) to convert the rest of the tangents to secants. Note that this method does require that we have at least one secant in the integral as well. If there aren't any secants then we'll need to do something different.

If the exponent on the secant is even and the exponent on the tangent is odd then we can use either case. Again, it will be easier to convert the term with the smallest exponent.

Let's take a look at a couple of examples.

Example 5 Evaluate the following integral.

$$\int \sec^9 x \tan^5 x\, dx$$

Solution

First note that since the exponent on the secant isn't even we can't use the substitution $u = \tan x$. However, the exponent on the tangent is odd and we've got a secant in the integral and so we will be able to use the substitution $u = \sec x$. This means stripping out a single tangent (along with a secant) and converting the remaining tangents to secants using (4).

Here's the work for this integral.

$$\begin{aligned}\int \sec^9 x \tan^5 x\, dx &= \int \sec^8 x \tan^4 x \tan x \sec x\, dx \\ &= \int \sec^8 x \left(\sec^2 x - 1\right)^2 \tan x \sec x\, dx \qquad\qquad u = \sec x \\ &= \int u^8 \left(u^2 - 1\right)^2 du \\ &= \int u^{12} - 2u^{10} + u^8\, du \\ &= \frac{1}{13}\sec^{13} x - \frac{2}{11}\sec^{11} x + \frac{1}{9}\sec^9 x + c\end{aligned}$$

Example 6 Evaluate the following integral.

$$\int \sec^4 x \tan^6 x\, dx$$

Solution

So, in this example the exponent on the tangent is even so the substitution $u = \sec x$ won't work. The exponent on the secant is even and so we can use the substitution $u = \tan x$ for this integral. That means that we need to strip out two secants and convert the rest to tangents. Here is the work for this integral.

$$\begin{aligned}\int \sec^4 x \tan^6 x\, dx &= \int \sec^2 x \tan^6 x \sec^2 x\, dx \\ &= \int \left(\tan^2 x + 1\right) \tan^6 x \sec^2 x\, dx \qquad\qquad u = \tan x \\ &= \int \left(u^2 + 1\right) u^6\, du \\ &= \int u^8 + u^6\, du \\ &= \frac{1}{9}\tan^9 x + \frac{1}{7}\tan^7 x + c\end{aligned}$$

Both of the previous examples fit very nicely into the patterns discussed above and so were not all that difficult to work. However, there are a couple of exceptions to the patterns above and in these cases there is no single method that will work for every problem. Each integral will be different and may require different solution methods in order to evaluate the integral.

Let's first take a look at a couple of integrals that have odd exponents on the tangents, but no secants. In these cases we can't use the substitution $u = \sec x$ since it requires there to be at least one secant in the integral.

Example 7 Evaluate the following integral.

$$\int \tan x \, dx$$

Solution

To do this integral all we need to do is recall the definition of tangent in terms of sine and cosine and then this integral is nothing more than a Calculus I substitution.

$$\begin{aligned} \int \tan x \, dx &= \int \frac{\sin x}{\cos x} dx & \qquad u = \cos x \\ &= -\int \frac{1}{u} du & \\ &= -\ln\left|\cos x\right| + c & \qquad r \ln x = \ln x^r \\ &= \ln\left|\cos x\right|^{-1} + c & \\ &\ln\left|\sec x\right| + c & \end{aligned}$$

Example 8 Evaluate the following integral.

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

Solution

The trick to this one is do the following manipulation of the integrand.

$$\begin{aligned} \int \tan^3 x \, dx &= \int \tan x \tan^2 x \, dx \\ &= \int \tan x \left(\sec^2 x - 1\right) dx \\ &= \int \tan x \sec^2 x \, dx - \int \tan x \, dx \end{aligned}$$

We can now use the substitution $u = \tan x$ on the first integral and the results from the previous example on the second integral.

The integral is then,

$$\int \tan^3 x \, dx = \frac{1}{2}\tan^2 x - \ln\left|\sec x\right| + c$$

Note that all odd powers of tangent (with the exception of the first power) can be integrated using the same method we used in the previous example. For instance,

$$\int \tan^5 x \, dx = \int \tan^3 x \left(\sec^2 x - 1\right) dx = \int \tan^3 x \sec^2 x \, dx - \int \tan^3 x \, dx$$

So, a quick substitution ($u = \tan x$) will give us the first integral and the second integral will always be the previous odd power.

Now let's take a look at a couple of examples in which the exponent on the secant is odd and the exponent on the tangent is even. In these cases the substitutions used above won't work.

It should also be noted that both of the following two integrals are integrals that we'll be seeing on occasion in later sections of this chapter and in later chapters. Because of this it wouldn't be a bad idea to make a note of these results so you'll have them ready when you need them later.

Example 9 Evaluate the following integral.

$$\int \sec x \, dx$$

Solution

This one isn't too bad once you see what you've got to do. By itself the integral can't be done. However, if we manipulate the integrand as follows we can do it.

$$\begin{aligned} \int \sec x \, dx &= \int \frac{\sec x \left(\sec x + \tan x\right)}{\sec x + \tan x} \, dx \\ &= \int \frac{\sec^2 x + \tan x \sec x}{\sec x + \tan x} \, dx \end{aligned}$$

In this form we can do the integral using the substitution $u = \sec x + \tan x$. Doing this gives,

$$\int \sec x \, dx = \ln\left|\sec x + \tan x\right| + c$$

The idea used in the above example is a nice idea to keep in mind. Multiplying the numerator and denominator of a term by the same term above can, on occasion, put the integral into a form that can be integrated. Note that this method won't always work and even when it does it won't always be clear what you need to multiply the numerator and denominator by. However, when it does work and you can figure out what term you need it can greatly simplify the integral.

Here's the next example.

Example 10 Evaluate the following integral.

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

Solution

This one is different from any of the other integrals that we've done in this section. The first step to doing this integral is to perform integration by parts using the following choices for *u* and *dv*.

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

Note that using integration by parts on this problem is not an obvious choice, but it does work very nicely here. After doing integration by parts we have,

$$\int \sec^3 x \, dx = \sec x \tan x - \int \sec x \tan^2 x \, dx$$

Now the new integral also has an odd exponent on the secant and an even exponent on the tangent and so the previous examples of products of secants and tangents still won't do us any good.

To do this integral we'll first write the tangents in the integral in terms of secants. Again, this is not necessarily an obvious choice but it's what we need to do in this case.

$$\begin{aligned} \int \sec^3 x \, dx &= \sec x \tan x - \int \sec x \left(\sec^2 x - 1\right) dx \\ &= \sec x \tan x - \int \sec^3 x \, dx + \int \sec x \, dx \end{aligned}$$

Now, we can use the results from the previous example to do the second integral and notice that the first integral is exactly the integral we're being asked to evaluate with a minus sign in front. So, add it to both sides to get,

$$2\int \sec^3 x\,dx = \sec x \tan x + \ln\left|\sec x + \tan x\right|$$

Finally divide by two and we're done.

$$\int \sec^3 x\,dx = \frac{1}{2}\left(\sec x \tan x + \ln\left|\sec x + \tan x\right|\right) + c$$

Again, note that we've again used the idea of integrating the right side until the original integral shows up and then moving this to the left side and dividing by its coefficient to complete the evaluation. We first saw this in the Integration by Parts section and noted at the time that this was a nice technique to remember. Here is another example of this technique.

Now that we've looked at products of secants and tangents let's also acknowledge that because we can relate cosecants and cotangents by

$$1 + \cot^2 x = \csc^2 x$$

all of the work that we did for products of secants and tangents will also work for products of cosecants and cotangents. We'll leave it to you to verify that.

There is one final topic to be discussed in this section before moving on.

To this point we've looked only at products of sines and cosines and products of secants and tangents. However, the methods used to do these integrals can also be used on some quotients involving sines and cosines and quotients involving secants and tangents (and hence quotients involving cosecants and cotangents).

Let's take a quick look at an example of this.

Example 11 Evaluate the following integral.

$$\int \frac{\sin^7 x}{\cos^4 x}\,dx$$

Solution

If this were a product of sines and cosines we would know what to do. We would strip out a sine (since the exponent on the sine is odd) and convert the rest of the sines to cosines. The same idea will work in this case. We'll strip out a sine from the numerator and convert the rest to cosines as follows,

$$\begin{aligned}\int \frac{\sin^7 x}{\cos^4 x}\,dx &= \int \frac{\sin^6 x}{\cos^4 x}\sin x\,dx \\ &= \int \frac{\left(\sin^2 x\right)^3}{\cos^4 x}\sin x\,dx \\ &= \int \frac{\left(1-\cos^2 x\right)^3}{\cos^4 x}\sin x\,dx\end{aligned}$$

At this point all we need to do is use the substitution $u = \cos x$ and we're done.

$$\begin{aligned}\int \frac{\sin^7 x}{\cos^4 x}\,dx &= -\int \frac{\left(1-u^2\right)^3}{u^4}\,du \\ &= -\int u^{-4} - 3u^{-2} + 3 - u^2\,du \\ &= -\left(-\frac{1}{3}\frac{1}{u^3} + 3\frac{1}{u} + 3u - \frac{1}{3}u^3\right) + c \\ &= \frac{1}{3\cos^3 x} - \frac{3}{\cos x} - 3\cos x + \frac{1}{3}\cos^3 x + c\end{aligned}$$

So, under the right circumstances, we can use the ideas developed to help us deal with products of trig functions to deal with quotients of trig functions. The natural question then, is just what are the right circumstances?

First notice that if the quotient had been reversed as in this integral,

$$\int \frac{\cos^4 x}{\sin^7 x}\,dx$$

we wouldn't have been able to strip out a sine.

$$\int \frac{\cos^4 x}{\sin^7 x}\,dx = \int \frac{\cos^4 x}{\sin^6 x}\,\frac{1}{\sin x}\,dx$$

In this case the "stripped out" sine remains in the denominator and it won't do us any good for the substitution $u = \cos x$ since this substitution requires a sine in the numerator of the quotient. Also note that, while we could convert the sines to cosines, the resulting integral would still be a fairly difficult integral.

So, we can use the methods we applied to products of trig functions to quotients of trig functions provided the term that needs parts stripped out in is the numerator of the quotient.

Trig Substitutions

As we have done in the last couple of sections, let's start off with a couple of integrals that we should already be able to do with a standard substitution.

$$\int x\sqrt{25x^2-4}\,dx = \frac{1}{75}\left(25x^2-4\right)^{\frac{3}{2}}+c \qquad\qquad \int \frac{x}{\sqrt{25x^2-4}}\,dx = \frac{1}{25}\sqrt{25x^2-4}+c$$

Both of these used the substitution $u = 25x^2 - 4$ and at this point should be pretty easy for you to do. However, let's take a look at the following integral.

Example 1 Evaluate the following integral.

$$\int \frac{\sqrt{25x^2-4}}{x}\,dx$$

Solution

In this case the substitution $u = 25x^2 - 4$ will not work and so we're going to have to do something different for this integral.

It would be nice if we could get rid of the square root somehow. The following substitution will do that for us.

$$x = \frac{2}{5}\sec\theta$$

Do not worry about where this came from at this point. As we work the problem you will see that it works and that if we have a similar type of square root in the problem we can always use a similar substitution.

Before we actually do the substitution however let's verify the claim that this will allow us to get rid of the square root.

$$\sqrt{25x^2-4} = \sqrt{25\left(\frac{4}{25}\right)\sec^2\theta-4} = \sqrt{4\left(\sec^2\theta-1\right)} = 2\sqrt{\sec^2\theta-1}$$

To get rid of the square root all we need to do is recall the relationship,

$$\tan^2\theta+1=\sec^2\theta \qquad \Rightarrow \qquad \sec^2\theta-1=\tan^2\theta$$

Using this fact the square root becomes,

$$\sqrt{25x^2-4} = 2\sqrt{\tan^2\theta} = 2\left|\tan\theta\right|$$

Note the presence of the absolute value bars there. These are important. Recall that

$$\sqrt{x^2} = \left|x\right|$$

There should always be absolute value bars at this stage. If we knew that $\tan\theta$ was always positive or always negative we could eliminate the absolute value bars using,

$$\left|x\right| = \begin{cases} x & \text{if } x \ge 0 \\ -x & \text{if } x < 0 \end{cases}$$

Without limits we won't be able to determine if $\tan\theta$ is positive or negative, however, we will need to eliminate them in order to do the integral. Therefore, since we are doing an indefinite integral we will assume that $\tan\theta$ will be positive and so we can drop the absolute value bars. This gives,

$$\sqrt{25x^2-4} = 2\tan\theta$$

So, we were able to eliminate the square root using this substitution. Let's now do the substitution and see what we get. In doing the substitution don't forget that we'll also need to substitute for the *dx*. This is easy enough to get from the substitution.

$$x = \frac{2}{5}\sec\theta \qquad \Rightarrow \qquad dx = \frac{2}{5}\sec\theta\tan\theta\,d\theta$$

Using this substitution the integral becomes,

$$\begin{aligned}\int \frac{\sqrt{25x^2-4}}{x}\,dx &= \int \frac{2\tan\theta}{\frac{2}{5}\sec\theta}\left(\frac{2}{5}\sec\theta\tan\theta\right)d\theta \\ &= 2\int \tan^2\theta\,d\theta\end{aligned}$$

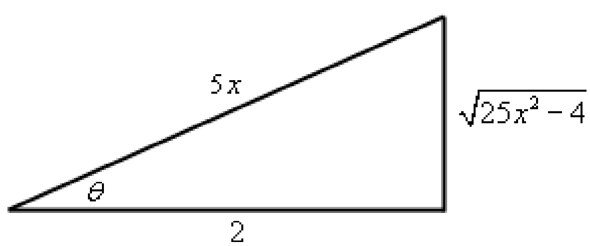
With this substitution we were able to reduce the given integral to an integral involving trig functions and we saw how to do these problems in the previous section. Let's finish the integral.

$$\begin{aligned}\int \frac{\sqrt{25x^2-4}}{x}\,dx &= 2\int \sec^2\theta - 1\,d\theta \\ &= 2\left(\tan\theta - \theta\right) + c\end{aligned}$$

So, we've got an answer for the integral. Unfortunately the answer isn't given in *x*'s as it should be. So, we need to write our answer in terms of *x*. We can do this with some right triangle trig. From our original substitution we have,

$$\sec\theta = \frac{5x}{2} = \frac{\text{hypotenuse}}{\text{adjacent}}$$

This gives the following right triangle.



From this we can see that,

$$\tan\theta = \frac{\sqrt{25x^2-4}}{2}$$

We can deal with the θ in one of any variety of ways. From our substitution we can see that,

$$\theta = \sec^{-1}\left(\frac{5x}{2}\right)$$

While this is a perfectly acceptable method of dealing with the θ we can use any of the possible six inverse trig functions and since sine and cosine are the two trig functions most people are familiar with we will usually use the inverse sine or inverse cosine. In this case we'll use the inverse cosine.

$$\theta = \cos^{-1}\left(\frac{2}{5x}\right)$$

So, with all of this the integral becomes,

$$\begin{aligned}\int \frac{\sqrt{25x^2-4}}{x}\,dx &= 2\left(\frac{\sqrt{25x^2-4}}{2} - \cos^{-1}\left(\frac{2}{5x}\right)\right) + c \\ &= \sqrt{25x^2-4} - 2\cos^{-1}\left(\frac{2}{5x}\right) + c\end{aligned}$$

We now have the answer back in terms of *x*.

Wow! That was a lot of work. Most of these won't take as long to work however. This first one needed lot's of explanation since it was the first one. The remaining examples won't need quite as much explanation and so won't take as long to work.

However, before we move onto more problems let's first address the issue of definite integrals and how the process differs in these cases.

Example 2 Evaluate the following integral.

$$\int_{\frac{2}{5}}^{\frac{4}{5}} \frac{\sqrt{25x^2-4}}{x}\,dx$$

Solution

The limits here won't change the substitution so that will remain the same.

$$x = \frac{2}{5}\sec\theta$$

Using this substitution the square root still reduces down to,

$$\sqrt{25x^2-4} = 2\left|\tan\theta\right|$$

However, unlike the previous example we can't just drop the absolute value bars. In this case we've got limits on the integral and so we can use the limits as well as the substitution to determine the range of θ that we're in. Once we've got that we can determine how to drop the absolute value bars.

Here's the limits of θ.

$$x = \frac{2}{5} \quad \Rightarrow \quad \frac{2}{5} = \frac{2}{5}\sec\theta \quad \Rightarrow \quad \theta = 0$$
$$x = \frac{4}{5} \quad \Rightarrow \quad \frac{4}{5} = \frac{2}{5}\sec\theta \quad \Rightarrow \quad \theta = \frac{\pi}{3}$$

So, if we are in the range $\frac{2}{5} \le x \le \frac{4}{5}$ then θ is in the range of $0 \le \theta \le \frac{\pi}{3}$ and in this range of θ's tangent is positive and so we can just drop the absolute value bars.

Let's do the substitution. Note that the work is identical to the previous example and so most of it is left out. We'll pick up at the final integral and then do the substitution.

$$\begin{aligned}\int_{\frac{2}{5}}^{\frac{4}{5}} \frac{\sqrt{25x^2-4}}{x}\,dx &= 2\int_0^{\frac{\pi}{3}} \sec^2\theta - 1\,d\theta \\ &= 2\left(\tan\theta - \theta\right)\Big|_0^{\pi/3} \\ &= 2\sqrt{3} - \frac{2\pi}{3}\end{aligned}$$

Note that because of the limits we didn't need to resort to a right triangle to complete the problem.

Let's take a look at a different set of limits for this integral.

Example 3 Evaluate the following integral.

$$\int_{-\frac{4}{5}}^{-\frac{2}{5}} \frac{\sqrt{25x^2-4}}{x}\,dx$$

Solution

Again, the substitution and square root are the same as the first two examples.

$$x = \frac{2}{5}\sec\theta \qquad\qquad \sqrt{25x^2-4} = 2\left|\tan\theta\right|$$

Let's next see the limits θ for this problem.

$$x = -\frac{2}{5} \quad \Rightarrow \quad -\frac{2}{5} = \frac{2}{5}\sec\theta \quad \Rightarrow \quad \theta = \pi$$
$$x = -\frac{4}{5} \quad \Rightarrow \quad -\frac{4}{5} = \frac{2}{5}\sec\theta \quad \Rightarrow \quad \theta = \frac{2\pi}{3}$$

Note that in determining the value of θ we used the smallest positive value. Now for this range of x's we have $\frac{2\pi}{3} \le \theta \le \pi$ and in this range of θ tangent is negative and so in this case we can drop the absolute value bars, but will need to add in a minus sign upon doing so. In other words,

$$\sqrt{25x^2-4} = -2\tan\theta$$

So, the only change this will make in the integration process is to put a minus sign in front of the integral. The integral is then,

$$\int_{-\frac{4}{5}}^{-\frac{2}{5}} \frac{\sqrt{25x^2-4}}{x}\,dx = -2\int_{\frac{2\pi}{3}}^{\pi} \sec^2\theta - 1\,d\theta$$
$$= -2\left(\tan\theta - \theta\right)\Big|_{2\pi/3}^{\pi}$$
$$= \frac{2\pi}{3} - 2\sqrt{3}$$

In the last two examples we saw that we have to be very careful with definite integrals. We need to make sure that we determine the limits on θ and whether or not this will mean that we can drop the absolute value bars or if we need to add in a minus sign when we drop them.

Before moving on to the next example let's get the general form for the substitution that we used in the previous set of examples.

$$\sqrt{b^2x^2 - a^2} \qquad \Rightarrow \qquad x = \frac{a}{b}\sec\theta$$

Let's work a new and different type of example.

Example 4 Evaluate the following integral.

$$\int \frac{1}{x^4\sqrt{9-x^2}}\,dx$$

Solution

Now, the square root in this problem looks to be (almost) the same as the previous ones so let's try the same type of substitution and see if it will work here as well.

$$x = 3\sec\theta$$

Using this substitution the square root becomes,

$$\sqrt{9-x^2} = \sqrt{9-9\sec^2\theta} = 3\sqrt{1-\sec^2\theta} = 3\sqrt{-\tan^2\theta}$$

So using this substitution we will end up with a negative quantity (the tangent squared is always positive of course) under the square root and this will be trouble. Using this substitution will give complex values and we don't want that. So, using secant for the substitution won't work.

However, the following substitution (and differential) will work.

$$x = 3\sin\theta \qquad\qquad dx = 3\cos\theta\,d\theta$$

With this substitution the square root is,

$$\sqrt{9-x^2} = 3\sqrt{1-\sin^2\theta} = 3\sqrt{\cos^2\theta} = 3\left|\cos\theta\right| = 3\cos\theta$$

We were able to drop the absolute value bars because we are doing an indefinite integral and so we'll assume that everything is positive.

The integral is now,

$$\int \frac{1}{x^4\sqrt{9-x^2}}\,dx = \int \frac{1}{81\sin^4\theta\left(3\cos\theta\right)} 3\cos\theta\,d\theta$$
$$= \frac{1}{81}\int \frac{1}{\sin^4\theta}\,d\theta$$
$$= \frac{1}{81}\int \csc^4\theta\,d\theta$$

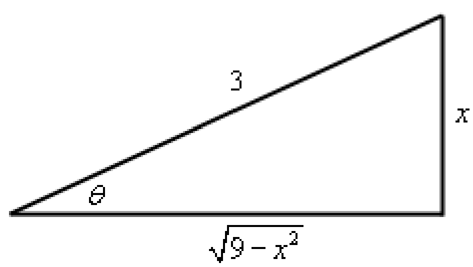
In the previous section we saw how to deal with integrals in which the exponent on the secant was even and since cosecants behave an awful lot like secants we should be able to do something similar with this.

Here is the integral.

$$\int \frac{1}{x^4\sqrt{9-x^2}}\,dx = \frac{1}{81}\int \csc^2\theta\csc^2\theta\,d\theta$$
$$= \frac{1}{81}\int \left(\cot^2\theta+1\right)\csc^2\theta\,d\theta \qquad\qquad u=\cot\theta$$
$$= -\frac{1}{81}\int u^2+1\,du$$
$$= -\frac{1}{81}\left(\frac{1}{3}\cot^3\theta+\cot\theta\right)+c$$

Now we need to go back to *x*'s using a right triangle. Here is the right triangle for this problem and trig functions for this problem.

$$\sin\theta = \frac{x}{3} \qquad\qquad \cot\theta = \frac{\sqrt{9-x^2}}{x}$$



The integral is then,

$$\int \frac{1}{x^4\sqrt{9-x^2}}\,dx = -\frac{1}{81}\left(\frac{1}{3}\left(\frac{\sqrt{9-x^2}}{x}\right)^3+\frac{\sqrt{9-x^2}}{x}\right)+c$$
$$= -\frac{\left(9-x^2\right)^{\frac{3}{2}}}{243x^3}-\frac{\sqrt{9-x^2}}{81x}+c$$

Here's the general form for this type of square root.

$$\sqrt{a^2 - b^2x^2} \qquad \Rightarrow \qquad x = \frac{a}{b}\sin\theta$$

There is one final case that we need to look at. The next integral will also contain something that we need to make sure we can deal with.

Example 5 Evaluate the following integral.

$$\int_0^{\frac{1}{6}} \frac{x^5}{\left(36x^2+1\right)^{\frac{3}{2}}}\,dx$$

Solution

First, notice that there really is a square root in this problem even though it isn't explicitly written out. To see the root let's rewrite things a little.

$$\left(36x^2+1\right)^{\frac{3}{2}} = \left(\left(36x^2+1\right)^{\frac{1}{2}}\right)^3 = \left(\sqrt{36x^2+1}\right)^3$$

This square root is not in the form we saw in the previous examples. Here we will use the substitution for this root.

$$x = \frac{1}{6}\tan\theta \qquad\qquad dx = \frac{1}{6}\sec^2\theta\,d\theta$$

With this substitution the denominator becomes,

$$\left(\sqrt{36x^2+1}\right)^3 = \left(\sqrt{\tan^2\theta+1}\right)^3 = \left(\sqrt{\sec^2\theta}\right)^3 = \left|\sec\theta\right|^3$$

Now, because we have limits we'll need to convert them to θ so we can determine how to drop the absolute value bars.

$$x = 0 \qquad \Rightarrow \qquad 0 = \frac{1}{6}\tan\theta \qquad \Rightarrow \qquad \theta = 0$$

$$x = \frac{1}{6} \qquad \Rightarrow \qquad \frac{1}{6} = \frac{1}{6}\tan\theta \qquad \Rightarrow \qquad \theta = \frac{\pi}{4}$$

In this range of θ secant is positive and so we can drop the absolute value bars.

Here is the integral,

$$\int_0^{\frac{1}{6}} \frac{x^5}{\left(36x^2+1\right)^{\frac{3}{2}}}\,dx = \int_0^{\frac{\pi}{4}} \frac{\frac{1}{7776}\tan^5\theta}{\sec^3\theta}\left(\frac{1}{6}\sec^2\theta\right)d\theta$$

$$= \frac{1}{46656}\int_0^{\frac{\pi}{4}} \frac{\tan^5\theta}{\sec\theta}\,d\theta$$

There are several ways to proceed from this point. Normally with an odd exponent on the tangent we would strip one of them out and convert to secants. However, that would require that we also have a secant in the numerator which we don't have. Therefore, it seems like the best way to do this one would be to convert the integrand to sines and cosines.

$$\int_0^{\frac{1}{6}} \frac{x^5}{\left(36x^2+1\right)^{\frac{3}{2}}}\,dx = \frac{1}{46656}\int_0^{\frac{\pi}{4}} \frac{\sin^5\theta}{\cos^4\theta}\,d\theta$$

$$= \frac{1}{46656}\int_0^{\frac{\pi}{4}} \frac{\left(1-\cos^2\theta\right)^2}{\cos^4\theta}\sin\theta\,d\theta$$

We can now use the substitution $u=\cos\theta$ and we might as well convert the limits as well.

$$\theta = 0 \qquad\qquad u=\cos 0 = 1$$

$$\theta = \frac{\pi}{4} \qquad\qquad u = \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}$$

The integral is then,

$$\int_0^{\frac{1}{6}} \frac{x^5}{\left(36x^2+1\right)^{\frac{3}{2}}}\,dx = -\frac{1}{46656}\int_1^{\frac{\sqrt{2}}{2}} u^{-4}-2u^{-2}+1\,du$$

$$= -\frac{1}{46656}\left(-\frac{1}{3u^3}+\frac{2}{u}+u\right)\Bigg|_1^{\frac{\sqrt{2}}{2}}$$

$$= \frac{1}{17496}-\frac{11\sqrt{2}}{279936}$$

The general form for this final type of square root is

$$\sqrt{a^2+b^2x^2} \qquad \Rightarrow \qquad x=\frac{a}{b}\tan\theta$$

We have a couple of final examples to work in this section. Not all trig substitutions will just jump right out at us. Sometimes we need to do a little work on the integrand first to get it into the correct form and that is the point of the remaining examples.

Example 6 Evaluate the following integral.

$$\int \frac{x}{\sqrt{2x^2-4x-7}}\,dx$$

Solution

In this case the quantity under the root doesn't obviously fit into any of the cases we looked at above and in fact isn't in the any of the forms we saw in the previous examples. Note however that if we complete the square on the quadratic we can make it look somewhat like the above integrals.

Remember that completing the square requires a coefficient of one in front of the x^2. Once we have that we take half the coefficient of the x, square it, and then add and subtract it to the quantity. Here is the completing the square for this problem.

$$2\left(x^2-2x-\frac{7}{2}\right)=2\left(x^2-2x+1-1-\frac{7}{2}\right)=2\left(\left(x-1\right)^2-\frac{9}{2}\right)=2\left(x-1\right)^2-9$$

So, the root becomes,

$$\sqrt{2x^2-4x-7}=\sqrt{2\left(x-1\right)^2-9}$$

This looks like a secant substitution except we don't just have an *x* that is squared. That is okay, it will work the same way.

$$x-1=\frac{3}{\sqrt{2}}\sec\theta \qquad x=1+\frac{3}{\sqrt{2}}\sec\theta \qquad dx=\frac{3}{\sqrt{2}}\sec\theta\tan\theta\,d\theta$$

Using this substitution the root reduces to,

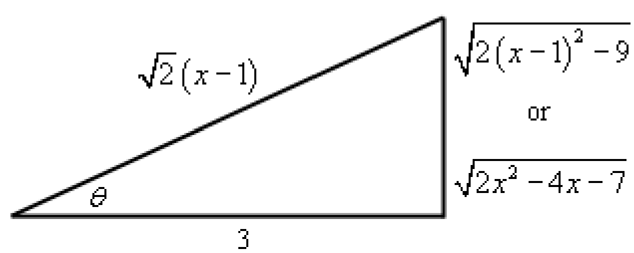
$$\sqrt{2x^2-4x-7}=\sqrt{2\left(x-1\right)^2-9}=\sqrt{9\sec^2\theta-9}=3\sqrt{\tan^2\theta}=3\left|\tan\theta\right|=3\tan\theta$$

Note we could drop the absolute value bars since we are doing an indefinite integral. Here is the integral.

$$\begin{aligned}\int\frac{x}{\sqrt{2x^2-4x-7}}\,dx&=\int\frac{1+\frac{3}{\sqrt{2}}\sec\theta}{3\tan\theta}\left(\frac{3}{\sqrt{2}}\sec\theta\tan\theta\right)d\theta\\&=\int\frac{1}{\sqrt{2}}\sec\theta+\frac{3}{2}\sec^2\theta\,d\theta\\&=\frac{1}{\sqrt{2}}\ln\left|\sec\theta+\tan\theta\right|+\frac{3}{2}\tan\theta+c\end{aligned}$$

And here is the right triangle for this problem.

$$\sec\theta=\frac{\sqrt{2}\left(x-1\right)}{3} \qquad \tan\theta=\frac{\sqrt{2x^2-4x-7}}{3}$$



The integral is then,

$$\int\frac{x}{\sqrt{2x^2-4x-7}}\,dx=\frac{1}{\sqrt{2}}\ln\left|\frac{\sqrt{2}\left(x-1\right)}{3}+\frac{\sqrt{2x^2-4x-7}}{3}\right|+\frac{\sqrt{2x^2-4x-7}}{2}+c$$

Example 7 Evaluate the following integral.

$$\int \mathbf{e}^{4x}\sqrt{1+\mathbf{e}^{2x}}\,dx$$

Solution

This doesn't look to be anything like the other problems in this section. However it is. To see this we first need to notice that,

$$\mathbf{e}^{2x} = \left(\mathbf{e}^{x}\right)^2$$

With this we can use the following substitution.

$$\mathbf{e}^{x} = \tan\theta \qquad\qquad \mathbf{e}^{x}\,dx = \sec^2\theta\,d\theta$$

Remember that to compute the differential all we do is differentiate both sides and then tack on dx or $d\theta$ onto the appropriate side.

With this substitution the square root becomes,

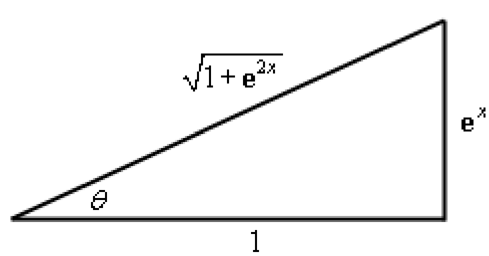
$$\sqrt{1+\mathbf{e}^{2x}} = \sqrt{1+\left(\mathbf{e}^{x}\right)^2} = \sqrt{1+\tan^2\theta} = \sqrt{\sec^2\theta} = \left|\sec\theta\right| = \sec\theta$$

Again, we can drop the absolute value bars because we are doing an indefinite integral. Here's the integral.

$$\begin{aligned}
\int \mathbf{e}^{4x}\sqrt{1+\mathbf{e}^{2x}}\,dx &= \int \mathbf{e}^{3x}\mathbf{e}^{x}\sqrt{1+\mathbf{e}^{2x}}\,dx \\
&= \int \left(\mathbf{e}^{x}\right)^3\sqrt{1+\mathbf{e}^{2x}}\left(\mathbf{e}^{x}\right)dx \\
&= \int \tan^3\theta\left(\sec\theta\right)\left(\sec^2\theta\right)d\theta \\
&= \int \left(\sec^2\theta-1\right)\sec^2\theta\sec\theta\tan\theta\,d\theta \qquad\qquad u=\sec\theta \\
&= \int u^4-u^2\,du \\
&= \frac{1}{5}\sec^5\theta-\frac{1}{3}\sec^3\theta+c
\end{aligned}$$

Here is the right triangle for this integral.

$$\tan\theta = \frac{\mathbf{e}^{x}}{1} \qquad\qquad \sec\theta = \frac{\sqrt{1+\mathbf{e}^{2x}}}{1} = \sqrt{1+\mathbf{e}^{2x}}$$



The integral is then,

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

So, as we've seen in the final two examples in this section some integrals that look nothing like the first few examples can in fact be turned into a trig substitution problem with a little work.

Before leaving this section let's summarize all three cases in one place.

$$\sqrt{a^2 - b^2x^2} \qquad \Rightarrow \qquad x = \frac{a}{b}\sin\theta$$

$$\sqrt{b^2x^2 - a^2} \qquad \Rightarrow \qquad x = \frac{a}{b}\sec\theta$$

$$\sqrt{a^2 + b^2x^2} \qquad \Rightarrow \qquad x = \frac{a}{b}\tan\theta$$

Partial Fractions

In this section we are going to take a look at integrals of rational expressions of polynomials and once again let's start this section out with an integral that we can already do so we can contrast it with the integrals that we'll be doing in this section.

$$\int \frac{2x-1}{x^2-x-6}\,dx = \int \frac{1}{u}\,du \qquad \text{using} \quad u = x^2 - x - 6 \quad \text{and} \quad du = \left(2x-1\right)dx$$
$$= \ln\left|x^2 - x - 6\right| + c$$

So, if the numerator is the derivative of the denominator (or a constant multiple of the derivative of the denominator) doing this kind of integral is fairly simple. However, often the numerator isn't the derivative of the denominator (or a constant multiple). For example, consider the following integral.

$$\int \frac{3x+11}{x^2-x-6}\,dx$$

In this case the numerator is definitely not the derivative of the denominator nor is it a constant multiple of the derivative of the denominator. Therefore, the simple substitution that we used above won't work. However, if we notice that the integrand can be broken up as follows,

$$\frac{3x+11}{x^2-x-6} = \frac{4}{x-3} - \frac{1}{x+2}$$

then the integral is actually quite simple.

$$\int \frac{3x+11}{x^2-x-6}\,dx = \int \frac{4}{x-3} - \frac{1}{x+2}\,dx$$
$$= 4\ln\left|x-3\right| - \ln\left|x+2\right| + c$$

This process of taking a rational expression and decomposing it into simpler rational expressions that we can add or subtract to get the original rational expression is called **partial fraction decomposition**. Many integrals involving rational expressions can be done if we first do partial fractions on the integrand.

So, let's do a quick review of partial fractions. We'll start with a rational expression in the form,

$$f\left(x\right) = \frac{P\left(x\right)}{Q\left(x\right)}$$

where both *P(x)* and *Q(x)* are polynomials and the degree of *P(x)* is smaller than the degree of *Q(x)*. Recall that the degree of a polynomial is the largest exponent in the polynomial. Partial fractions can only be done if the degree of the numerator is strictly less than the degree of the denominator. That is important to remember.

So, once we've determined that partial fractions can be done we factor the denominator as completely as possible. Then for each factor in the denominator we can use the following table to determine the term(s) we pick up in the partial fraction decomposition.

Factor in denominator	Term in partial fraction decomposition
$ax+b$	$\dfrac{A}{ax+b}$
$\left(ax+b\right)^k$	$\dfrac{A_1}{ax+b}+\dfrac{A_2}{\left(ax+b\right)^2}+\cdots+\dfrac{A_k}{\left(ax+b\right)^k}\ ,\quad k=1,2,3,\ldots$
ax^2+bx+c	$\dfrac{Ax+B}{ax^2+bx+c}$
$\left(ax^2+bx+c\right)^k$	$\dfrac{A_1x+B_1}{ax^2+bx+c}+\dfrac{A_2x+B_2}{\left(ax^2+bx+c\right)^2}+\cdots+\dfrac{A_kx+B_k}{\left(ax^2+bx+c\right)^k}\ ,\quad k=1,2,3,\ldots$

Notice that the first and third cases are really special cases of the second and fourth cases respectively.

There are several methods for determining the coefficients for each term and we will go over each of those in the following examples.

Let's start the examples by doing the integral above.

Example 1 Evaluate the following integral.

$$\int \frac{3x+11}{x^2-x-6}\,dx$$

Solution

The first step is to factor the denominator as much as possible and get the form of the partial fraction decomposition. Doing this gives,

$$\frac{3x+11}{\left(x-3\right)\left(x+2\right)}=\frac{A}{x-3}+\frac{B}{x+2}$$

The next step is to actually add the right side back up.

$$\frac{3x+11}{\left(x-3\right)\left(x+2\right)}=\frac{A\left(x+2\right)+B\left(x-3\right)}{\left(x-3\right)\left(x+2\right)}$$

Now, we need to choose A and B so that the numerators of these two are equal for every x. To do this we'll need to set the numerators equal.

$$3x+11=A\left(x+2\right)+B\left(x-3\right)$$

Note that in most problems we will go straight from the general form of the decomposition to this step and not bother with actually adding the terms back up. The only point to adding the terms is to get the numerator and we can get that without actually writing down the results of the addition.

At this point we have one of two ways to proceed. One way will always work, but is often more work. The other, while it won't always work, is often quicker when it does work. In this case both will work and so we'll use the quicker way for this example. We'll take a look at the other method in a later example.

What we're going to do here is to notice that the numerators must be equal for *any x* that we would choose to use. In particular the numerators must be equal for $x=-2$ and $x=3$. So, let's plug these in and see what we get.

$$\begin{array}{llll} x=-2 & 5=A(0)+B(-5) & \Rightarrow & B=-1 \\ x=3 & 20=A(5)+B(0) & \Rightarrow & A=4 \end{array}$$

So, by carefully picking the *x*'s we got the unknown constants to quickly drop out. Note that these are the values we claimed they would be above.

At this point there really isn't a whole lot to do other than the integral.

$$\begin{aligned} \int \frac{3x+11}{x^2-x-6}\,dx &= \int \frac{4}{x-3}-\frac{1}{x+2}\,dx \\ &= \int \frac{4}{x-3}\,dx-\int \frac{1}{x+2}\,dx \\ &= 4\ln|x-3|-\ln|x+2|+c \end{aligned}$$

Recall that to do this integral we first split it up into two integrals and then used the substitutions,

$$u=x-3 \qquad\qquad v=x+2$$

on the integrals to get the final answer.

Before moving onto the next example a couple of quick notes are in order here. First, many of the integrals in partial fractions problems come down to the type of integral seen above. Make sure that you can do those integrals.

There is also another integral that often shows up in these kinds of problems so we may as well give the formula for it here since we are already on the subject.

$$\int \frac{1}{x^2+a^2}\,dx=\frac{1}{a}\tan^{-1}\left(\frac{x}{a}\right)+c$$

It will be an example or two before we use this so don't forget about it.

Now, let's work some more examples.

Example 2 Evaluate the following integral.

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

Solution

We won't be putting as much detail into this solution as we did in the previous example. The first thing is to factor the denominator and get the form of the partial fraction decomposition.

$$\frac{x^2+4}{x(x+2)(3x-2)}=\frac{A}{x}+\frac{B}{x+2}+\frac{C}{3x-2}$$

The next step is to set numerators equal. If you need to actually add the right side together to get the numerator for that side then you should do so, however, it will definitely make the problem quicker if you can do the addition in your head to get,

$$x^2+4=A(x+2)(3x-2)+Bx(3x-2)+Cx(x+2)$$

As with the previous example it looks like we can just pick a few values of x and find the constants so let's do that.

$$\begin{aligned} x&=0 & 4&=A(2)(-2) & &\Rightarrow & A&=-1 \\ x&=-2 & 8&=B(-2)(-8) & &\Rightarrow & B&=\frac{1}{2} \\ x&=\frac{2}{3} & \frac{40}{9}&=C\left(\frac{2}{3}\right)\left(\frac{8}{3}\right) & &\Rightarrow & C&=\frac{40}{16}=\frac{5}{2} \end{aligned}$$

Note that unlike the first example most of the coefficients here are fractions. That is not unusual so don't get excited about it when it happens.

Now, let's do the integral.

$$\begin{aligned} \int \frac{x^2+4}{3x^3+4x^2-4x}\,dx &= \int -\frac{1}{x}+\frac{\frac{1}{2}}{x+2}+\frac{\frac{5}{2}}{3x-2}\,dx \\ &= -\ln|x|+\frac{1}{2}\ln|x+2|+\frac{5}{6}\ln|3x-2|+c \end{aligned}$$

Again, as noted above, integrals that generate natural logarithms are very common in these problems so make sure you can do them.

Example 3 Evaluate the following integral.

$$\int \frac{x^2-29x+5}{(x-4)^2\left(x^2+3\right)}\,dx$$

Solution

This time the denominator is already factored so let's just jump right to the partial fraction decomposition.

$$\frac{x^2-29x+5}{(x-4)^2\left(x^2+3\right)}=\frac{A}{x-4}+\frac{B}{(x-4)^2}+\frac{Cx+D}{x^2+3}$$

Setting numerators gives,

$$x^2-29x+5=A(x-4)\left(x^2+3\right)+B\left(x^2+3\right)+(Cx+D)(x-4)^2$$

In this case we aren't going to be able to just pick values of x that will give us all the constants. Therefore, we will need to work this the second (and often longer) way. The first step is to multiply out the right side and collect all the like terms together. Doing this gives,

$$x^2-29x+5=(A+C)x^3+(-4A+B-8C+D)x^2+(3A+16C-8D)x-12A+3B+16D$$

Now we need to choose A, B, C, and D so that these two are equal. In other words we will need to set the coefficients of like powers of x equal. This will give a system of equations that can be solved.

$$\left.\begin{aligned} x^3: &\quad A+C=0 \\ x^2: &\quad -4A+B-8C+D=1 \\ x^1: &\quad 3A+16C-8D=-29 \\ x^0: &\quad -12A+3B+16D=5 \end{aligned}\right\} \quad \Rightarrow \quad A=1, B=-5, C=-1, D=2$$

Note that we used x^0 to represent the constants. Also note that these systems can often be quite large and have a fair amount of work involved in solving them. The best way to deal with these is to use some form of computer aided solving techniques.

Now, let's take a look at the integral.

$$\begin{aligned} \int \frac{x^2-29x+5}{(x-4)^2(x^2+3)}\,dx &= \int \frac{1}{x-4} - \frac{5}{(x-4)^2} + \frac{-x+2}{x^2+3}\,dx \\ &= \int \frac{1}{x-4} - \frac{5}{(x-4)^2} - \frac{x}{x^2+3} + \frac{2}{x^2+3}\,dx \\ &= \ln|x-4| + \frac{5}{x-4} - \frac{1}{2}\ln|x^2+3| + \frac{2}{\sqrt{3}}\tan^{-1}\left(\frac{x}{\sqrt{3}}\right) + c \end{aligned}$$

In order to take care of the third term we needed to split it up into two separate terms. Once we've done this we can do all the integrals in the problem. The first two use the substitution $u = x-4$, the third uses the substitution $v = x^2+3$ and the fourth term uses the formula given above for inverse tangents.

Example 4 Evaluate the following integral.

$$\int \frac{x^3+10x^2+3x+36}{(x-1)(x^2+4)^2}\,dx$$

Solution

Let's first get the general form of the partial fraction decomposition.

$$\frac{x^3+10x^2+3x+36}{(x-1)(x^2+4)^2} = \frac{A}{x-1} + \frac{Bx+C}{x^2+4} + \frac{Dx+E}{(x^2+4)^2}$$

Now, set numerators equal, expand the right side and collect like terms.

$$\begin{aligned} x^3+10x^2+3x+36 &= A(x^2+4)^2 + (Bx+C)(x-1)(x^2+4) + (Dx+E)(x-1) \\ &= (A+B)x^4 + (C-B)x^3 + (8A+4B-C+D)x^2 + \\ &\qquad (-4B+4C-D+E)x + 16A - 4C - E \end{aligned}$$

Setting coefficient equal gives the following system.

$$\left.\begin{aligned} x^4: &\quad A+B=0 \\ x^3: &\quad C-B=1 \\ x^2: &\quad 8A+4B-C+D=10 \\ x^1: &\quad -4B+4C-D+E=3 \\ x^0: &\quad 16A-4C-E=36 \end{aligned}\right\} \quad \Rightarrow \quad A=2, B=-2, C=-1, D=1, E=0$$

Don't get excited if some of the coefficients end up being zero. It happens on occasion.

Here's the integral.

$$\begin{aligned} \int \frac{x^3+10x^2+3x+36}{(x-1)\left(x^2+4\right)^2}\,dx &= \int \frac{2}{x-1} + \frac{-2x-1}{x^2+4} + \frac{x}{\left(x^2+4\right)^2}\,dx \\ &= \int \frac{2}{x-1} - \frac{2x}{x^2+4} - \frac{1}{x^2+4} + \frac{x}{\left(x^2+4\right)^2}\,dx \\ &= 2\ln|x-1| - \ln\left|x^2+4\right| - \frac{1}{2}\tan^{-1}\left(\frac{x}{2}\right) - \frac{1}{2}\frac{1}{x^2+4} + c \end{aligned}$$

To this point we've only looked at rational expressions where the degree of the numerator was strictly less that the degree of the denominator. Of course not all rational expressions will fit into this form and so we need to take a look at a couple of examples where this isn't the case.

Example 5 Evaluate the following integral.

$$\int \frac{x^4-5x^3+6x^2-18}{x^3-3x^2}\,dx$$

Solution

So, in this case the degree of the numerator is 4 and the degree of the denominator is 3. Therefore, partial fractions can't be done on this rational expression.

To fix this up we'll need to do long division on this to get it into a form that we can deal with. Here is the work for that.

$$\begin{array}{r} x-2 \\ x^3-3x^2 \overline{\smash{\big)}\, x^4-5x^3+6x^2-18} \\ \underline{-\left(x^4-3x^3\right) \qquad\qquad\quad} \\ -2x^3+6x^2-18 \\ \underline{-\left(-2x^3+6x^2\right) \quad\;\;} \\ -18 \end{array}$$

So, from the long division we see that,

$$\frac{x^4-5x^3+6x^2-18}{x^3-3x^2}=x-2-\frac{18}{x^3-3x^2}$$

and the integral becomes,

$$\begin{aligned}\int \frac{x^4-5x^3+6x^2-18}{x^3-3x^2}\,dx &= \int x-2-\frac{18}{x^3-3x^2}\,dx\\ &= \int x-2\,dx-\int \frac{18}{x^3-3x^2}\,dx\end{aligned}$$

The first integral we can do easily enough and the second integral is now in a form that allows us to do partial fractions. So, let's get the general form of the partial fractions for the second integrand.

$$\frac{18}{x^2\left(x-3\right)}=\frac{A}{x}+\frac{B}{x^2}+\frac{C}{x-3}$$

Setting numerators equal gives us,

$$18=Ax\left(x-3\right)+B\left(x-3\right)+Cx^2$$

Now, there is a variation of the method we used in the first couple of examples that will work here. There are a couple of values of *x* that will allow us to quickly get two of the three constants, but there is no value of *x* that will just hand us the third.

What we'll do in this example is pick *x*'s to get the two constants that we can easily get and then we'll just pick another value of *x* that will be easy to work with (*i.e.* it won't give large/messy numbers anywhere) and then we'll use the fact that we also know the other two constants to find the third.

$$\begin{array}{lllll} x=0 & & 18=B\left(-3\right) & \Rightarrow & B=-6\\ x=3 & & 18=C\left(9\right) & \Rightarrow & C=2\\ x=1 & & 18=A\left(-2\right)+B\left(-2\right)+C=-2A+14 & \Rightarrow & A=-2\end{array}$$

The integral is then,

$$\begin{aligned}\int \frac{x^4-5x^3+6x^2-18}{x^3-3x^2}\,dx &= \int x-2\,dx-\int -\frac{2}{x}-\frac{6}{x^2}+\frac{2}{x-3}\,dx\\ &= \frac{1}{2}x^2-2x+2\ln\left|x\right|-\frac{6}{x}-2\ln\left|x-3\right|+c\end{aligned}$$

In the previous example there were actually two different ways of dealing with the x^2 in the denominator. One is to treat it as a quadratic which would give the following term in the decomposition

$$\frac{Ax+B}{x^2}$$

and the other is to treat it as a linear term in the following way,

$$x^2=\left(x-0\right)^2$$

which gives the following two terms in the decomposition,

$$\frac{A}{x}+\frac{B}{x^2}$$

We used the second way of thinking about it in our example. Notice however that the two will give identical partial fraction decompositions. So, why talk about this? Simple. This will work for x^2, but what about x^3 or x^4? In these cases we really will need to use the second way of thinking about these kinds of terms.

$$x^3 \;\Rightarrow\; \frac{A}{x}+\frac{B}{x^2}+\frac{C}{x^3} \qquad\qquad x^4 \;\Rightarrow\; \frac{A}{x}+\frac{B}{x^2}+\frac{C}{x^3}+\frac{D}{x^4}$$

Let's take a look at one more example.

Example 6 Evaluate the following integral.

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

Solution

In this case the numerator and denominator have the same degree. As with the last example we'll need to do long division to get this into the correct form. I'll leave the details of that to you to check.

$$\int \frac{x^2}{x^2-1}\,dx = \int 1+\frac{1}{x^2-1}\,dx = \int dx + \int \frac{1}{x^2-1}\,dx$$

So, we'll need to partial fraction the second integral. Here's the decomposition.

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

Setting numerator equal gives,

$$1 = A(x+1)+B(x-1)$$

Picking value of x gives us the following coefficients.

$$x=-1 \qquad 1=B(-2) \qquad \Rightarrow \qquad B=-\frac{1}{2}$$

$$x=1 \qquad 1=A(2) \qquad \Rightarrow \qquad A=\frac{1}{2}$$

The integral is then,

$$\begin{aligned}\int \frac{x^2}{x^2-1}\,dx &= \int dx + \int \frac{\frac{1}{2}}{x-1}-\frac{\frac{1}{2}}{x+1}\,dx \\ &= x+\frac{1}{2}\ln|x-1|-\frac{1}{2}\ln|x+1|+c\end{aligned}$$

Integrals Involving Roots

In this section we're going to look at an integration technique that can be useful for *some* integrals with roots in them. We've already seen some integrals with roots in them. Some can be done quickly with a simple Calculus I substitution and some can be done with trig substitutions.

However, not all integrals with roots will allow us to use one of these methods. Let's look at a couple of examples to see another technique that can be used on occasion to help with these integrals.

Example 1 Evaluate the following integral.

$$\int \frac{x+2}{\sqrt[3]{x-3}}\,dx$$

Solution

Sometimes when faced with an integral that contains a root we can use the following substitution to simplify the integral into a form that can be easily worked with.

$$u = \sqrt[3]{x-3}$$

So, instead of letting *u* be the stuff under the radical as we often did in Calculus I we let *u* be the whole radical. Now, there will be a little more work here since we will also need to know what *x* is so we can substitute in for that in the numerator and so we can compute the differential, *dx*. This is easy enough to get however. Just solve the substitution for *x* as follows,

$$x = u^3 + 3 \qquad\qquad dx = 3u^2\,du$$

Using this substitution the integral is now,

$$\begin{aligned}\int \frac{\left(u^3+3\right)+2}{u}3u^2 du &= \int 3u^4 + 15u\,du\\ &= \frac{3}{5}u^5 + \frac{15}{2}u^2 + c\\ &= \frac{3}{5}\left(x-3\right)^{\frac{5}{3}} + \frac{15}{2}\left(x-3\right)^{\frac{2}{3}} + c\end{aligned}$$

So, sometimes, when an integral contains the root $\sqrt[n]{g\left(x\right)}$ the substitution,

$$u = \sqrt[n]{g\left(x\right)}$$

can be used to simplify the integral into a form that we can deal with.

Let's take a look at another example real quick.

Example 2 Evaluate the following integral.

$$\int \frac{2}{x-3\sqrt{x+10}}\,dx$$

Solution

We'll do the same thing we did in the previous example. Here's the substitution and the extra work we'll need to do to get *x* in terms of *u*.

$$u = \sqrt{x+10} \qquad\qquad x = u^2 - 10 \qquad\qquad dx = 2u\,du$$

With this substitution the integral is,

$$\int \frac{2}{x - 3\sqrt{x+10}}\,dx = \int \frac{2}{u^2 - 10 - 3u}\left(2u\right)du = \int \frac{4u}{u^2 - 3u - 10}\,du$$

This integral can now be done with partial fractions.

$$\frac{4u}{\left(u-5\right)\left(u+2\right)} = \frac{A}{u-5} + \frac{B}{u+2}$$

Setting numerators equal gives,

$$4u = A\left(u+2\right) + B\left(u-5\right)$$

Picking value of u gives the coefficients.

$$u = -2 \qquad\qquad -8 = B\left(-7\right) \qquad\qquad B = \frac{8}{7}$$

$$u = 5 \qquad\qquad 20 = A\left(7\right) \qquad\qquad A = \frac{20}{7}$$

The integral is then,

$$\begin{aligned}\int \frac{2}{x - 3\sqrt{x+10}}\,dx &= \int \frac{\frac{20}{7}}{u-5} + \frac{\frac{8}{7}}{u+2}\,du \\ &= \frac{20}{7}\ln\left|u-5\right| + \frac{8}{7}\ln\left|u+2\right| + c \\ &= \frac{20}{7}\ln\left|\sqrt{x+10} - 5\right| + \frac{8}{7}\ln\left|\sqrt{x+10} + 2\right| + c\end{aligned}$$

So, we’ve seen a nice method to eliminate roots from the integral and put it into a form that we can deal with. Note however, that this won’t always work and sometimes the new integral will be just as difficult to do.

Integrals Involving Quadratics

To this point we've seen quite a few integrals that involve quadratics. A couple of examples are,

$$\int \frac{x}{x^2 \pm a}\,dx = \frac{1}{2}\ln\left|x^2 \pm a\right| + c \qquad\qquad \int \frac{1}{x^2+a^2}\,dx = \frac{1}{a}\tan^{-1}\left(\frac{x}{a}\right)$$

We also saw that integrals involving $\sqrt{b^2x^2-a^2}$, $\sqrt{a^2-b^2x^2}$ and $\sqrt{a^2+b^2x^2}$ could be done with a trig substitution.

Notice however that all of these integrals were missing an x term. They all consist of a quadratic term and a constant.

Some integrals involving general quadratics are easy enough to do. For instance, the following integral can be done with a quick substitution.

$$\begin{aligned}\int \frac{2x+3}{4x^2+12x-1}\,dx &= \frac{1}{4}\int \frac{1}{u}\,du \qquad\qquad \left(u = 4x^2+12x-1 \qquad du = 4\left(2x+3\right)dx\right)\\ &= \frac{1}{4}\ln\left|4x^2+12x-1\right| + c\end{aligned}$$

Some integrals with quadratics can be done with partial fractions. For instance,

$$\int \frac{10x-6}{3x^2+16x+5}\,dx = \int \frac{4}{x+5} - \frac{2}{3x+1}\,dx = 4\ln\left|x+5\right| - \frac{2}{3}\ln\left|3x+1\right| + c$$

Unfortunately, these methods won't work on a lot of integrals. A simple substitution will only work if the numerator is a constant multiple of the derivative of the denominator and partial fractions will only work if the denominator can be factored.

This section is how to deal with integrals involving quadratics when the techniques that we've looked at to this point simply won't work.

Back in the Trig Substitution section we saw how to deal with square roots that had a general quadratic in them. Let's take a quick look at another one like that since the idea involved in doing that kind of integral is exactly what we are going to need for the other integrals in this section.

Example 1 Evaluate the following integral.

$$\int \sqrt{x^2+4x+5}\,dx$$

Solution

Recall from the Trig Substitution section that in order to do a trig substitution here we first needed to complete the square on the quadratic. This gives,

$$x^2+4x+5 = x^2+4x+4-4+5 = \left(x+2\right)^2+1$$

After completing the square the integral becomes,

$$\int \sqrt{x^2+4x+5}\,dx = \int \sqrt{\left(x+2\right)^2+1}\,dx$$

Upon doing this we can identify the trig substitution that we need. Here it is,

$$x+2=\tan\theta \qquad x=\tan\theta-2 \qquad dx=\sec^2\theta\,d\theta$$

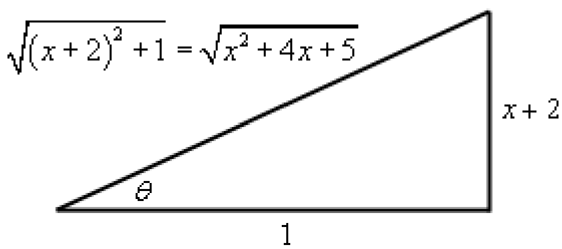
$$\sqrt{(x+2)^2+1}=\sqrt{\tan^2\theta+1}=\sqrt{\sec^2\theta}=\left|\sec\theta\right|=\sec\theta$$

Recall that since we are doing an indefinite integral we can drop the absolute value bars. Using this substitution the integral becomes,

$$\begin{aligned}\int\sqrt{x^2+4x+5}\,dx&=\int\sec^3\theta\,d\theta\\&=\frac{1}{2}\left(\sec\theta\tan\theta+\ln\left|\sec\theta+\tan\theta\right|\right)+c\end{aligned}$$

We can finish the integral out with the following right triangle.

$$\tan\theta=\frac{x+2}{1} \qquad \sec\theta=\frac{\sqrt{x^2+4x+5}}{1}=\sqrt{x^2+4x+5}$$



$$\int\sqrt{x^2+4x+5}\,dx=\frac{1}{2}\left((x+2)\sqrt{x^2+4x+5}+\ln\left|x+2+\sqrt{x^2+4x+5}\right|\right)+c$$

So, by completing the square we were able to take an integral that had a general quadratic in it and convert it into a form that allowed us to use a known integration technique.

Let's do a quick review of completing the square before proceeding. Here is the general completing the square formula that we'll use.

$$x^2+bx+c=x^2+bx+\left(\frac{b}{2}\right)^2-\left(\frac{b}{2}\right)^2+c=\left(x+\frac{b}{2}\right)^2+c-\frac{b^2}{4}$$

This will always take a general quadratic and write it in terms of a squared term and a constant term.

Recall as well that in order to do this we must have a coefficient of one in front of the x^2. If not we'll need to factor out the coefficient before completing the square. In other words,

$$ax^2+bx+c=a\left(\underbrace{x^2+\frac{b}{a}x+\frac{c}{a}}_{\text{complete the square on this!}}\right)$$

Now, let's see how completing the square can be used to do integrals that we aren't able to do at this point.

Example 2 Evaluate the following integral.

$$\int \frac{1}{2x^2 - 3x + 2}\,dx$$

Solution

Okay, this doesn't factor so partial fractions just won't work on this. Likewise, since the numerator is just "1" we can't use the substitution $u = 2x^2 - 3x + 8$. So, let's see what happens if we complete the square on the denominator.

$$\begin{aligned} 2x^2 - 3x + 2 &= 2\left(x^2 - \frac{3}{2}x + 1\right) \\ &= 2\left(x^2 - \frac{3}{2}x + \frac{9}{16} - \frac{9}{16} + 1\right) \\ & 2\left(\left(x - \frac{3}{4}\right)^2 + \frac{7}{16}\right) \end{aligned}$$

With this the integral is,

$$\int \frac{1}{2x^2 - 3x + 2}\,dx = \frac{1}{2}\int \frac{1}{\left(x - \frac{3}{4}\right)^2 + \frac{7}{16}}\,dx$$

Now this may not seem like all that great of a change. However, notice that we can now use the following substitution.

$$u = x - \frac{3}{4} \qquad\qquad du = dx$$

and the integral is now,

$$\int \frac{1}{2x^2 - 3x + 2}\,dx = \frac{1}{2}\int \frac{1}{u^2 + \frac{7}{16}}\,du$$

We can now see that this is an inverse tangent! So, using the formula from above we get,

$$\begin{aligned} \int \frac{1}{2x^2 - 3x + 2}\,dx &= \frac{1}{2}\left(\frac{4}{\sqrt{7}}\right)\tan^{-1}\left(\frac{4u}{\sqrt{7}}\right) + c \\ &= \frac{2}{\sqrt{7}}\tan^{-1}\left(\frac{4x-3}{\sqrt{7}}\right) + c \end{aligned}$$

Example 3 Evaluate the following integral.

$$\int \frac{3x - 1}{x^2 + 10x + 28}\,dx$$

Solution

This example is a little different from the previous one. In this case we do have an *x* in the numerator however the numerator still isn't a multiple of the derivative of the denominator and so a simple Calculus I substitution won't work.

So, let's again complete the square on the denominator and see what we get,

$$x^2 + 10x + 28 = x^2 + 10x + 25 - 25 + 28 = (x+5)^2 + 3$$

Upon completing the square the integral becomes,

$$\int \frac{3x-1}{x^2+10x+28}\,dx = \int \frac{3x-1}{(x+5)^2+3}\,dx$$

At this point we can use the same type of substitution that we did in the previous example. The only real difference is that we'll need to make sure that we plug the substitution back into the numerator as well.

$$u = x+5 \qquad\qquad x = u-5 \qquad\qquad dx = du$$

$$\begin{aligned}
\int \frac{3x-1}{x^2+10x+28}\,dx &= \int \frac{3(u-5)-1}{u^2+3}\,du \\
&= \int \frac{3u}{u^2+3} - \frac{16}{u^2+3}\,du \\
&= \frac{3}{2}\ln\left|u^2+3\right| - \frac{16}{\sqrt{3}}\tan^{-1}\left(\frac{u}{\sqrt{3}}\right) + c \\
&= \frac{3}{2}\ln\left|(x+5)^2+3\right| - \frac{16}{\sqrt{3}}\tan^{-1}\left(\frac{x+5}{\sqrt{3}}\right) + c
\end{aligned}$$

So, in general when dealing with an integral in the form,

$$\int \frac{Ax+B}{ax^2+bx+c}\,dx \tag{1}$$

Here we are going to assume that the denominator doesn't factor and the numerator isn't a constant multiple of the derivative of the denominator. In these cases we complete the square on the denominator and then do a substitution that will yield an inverse tangent and/or a logarithm depending on the exact form of the numerator.

Let's now take a look at a couple of integrals that are in the same general form as (1) except the denominator will also be raised to a power. In other words, let's look at integrals in the form,

$$\int \frac{Ax+B}{\left(ax^2+bx+c\right)^n}\,dx \tag{2}$$

Example 4 Evaluate the following integral.

$$\int \frac{x}{\left(x^2-6x+11\right)^3}\,dx$$

Solution

For the most part this integral will work the same as the previous two with one exception that will occur down the road. So, let's start by completing the square on the quadratic in the denominator.

$$x^2 - 6x + 11 = x^2 - 6x + 9 - 9 + 11 = (x-3)^2 + 2$$

The integral is then,

$$\int \frac{x}{\left(x^2 - 6x + 11\right)^3}\,dx = \int \frac{x}{\left[(x-3)^2 + 2\right]^3}\,dx$$

Now, we will use the same substitution that we've used to this point in the previous two examples.

$$u = x - 3 \qquad\qquad x = u + 3 \qquad\qquad dx = du$$

$$\begin{aligned}\int \frac{x}{\left(x^2 - 6x + 11\right)^3}\,dx &= \int \frac{u+3}{\left(u^2+2\right)^3}\,du \\ &= \int \frac{u}{\left(u^2+2\right)^3}\,du + \int \frac{3}{\left(u^2+2\right)^3}\,du\end{aligned}$$

Now, here is where the differences start cropping up. The first integral can be done with the substitution $v = u^2 + 2$ and isn't too difficult. The second integral however, can't be done with the substitution used on the first integral and it isn't an inverse tangent.

It turns out that a trig substitution will work nicely on the second integral and it will be the same as we did when we had square roots in the problem.

$$u = \sqrt{2}\tan\theta \qquad\qquad du = \sqrt{2}\sec^2\theta\,d\theta$$

With these two substitutions the integrals become,

$$\begin{aligned}\int \frac{x}{\left(x^2 - 6x + 11\right)^3}\,dx &= \frac{1}{2}\int \frac{1}{v^3}\,dv + \int \frac{3}{\left(2\tan^2\theta + 2\right)^3}\left(\sqrt{2}\sec^2\theta\right)d\theta \\ &= -\frac{1}{4}\frac{1}{v^2} + \int \frac{3\sqrt{2}\sec^2\theta}{8\left(\tan^2\theta + 1\right)^3}\,d\theta \\ &= -\frac{1}{4}\frac{1}{\left(u^2+2\right)^2} + \frac{3\sqrt{2}}{8}\int \frac{\sec^2\theta}{\left(\sec^2\theta\right)^3}\,d\theta \\ &= -\frac{1}{4}\frac{1}{\left((x-3)^2+2\right)^2} + \frac{3\sqrt{2}}{8}\int \frac{1}{\sec^4\theta}\,d\theta \\ &= -\frac{1}{4}\frac{1}{\left((x-3)^2+2\right)^2} + \frac{3\sqrt{2}}{8}\int \cos^4\theta\,d\theta\end{aligned}$$

Okay, at this point we've got two options for the remaining integral. We can either use the ideas we learned in the section about integrals involving trig integrals or we could use the following formula.

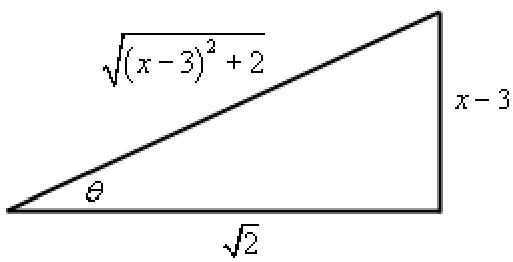
$$\int \cos^m \theta \, d\theta = \frac{1}{m} \sin\theta \cos^{m-1}\theta + \frac{m-1}{m}\int \cos^{m-2}\theta \, d\theta$$

Let's use this formula to do the integral.

$$\begin{aligned} \int \cos^4 \theta \, d\theta &= \frac{1}{4}\sin\theta\cos^3\theta + \frac{3}{4}\int \cos^2\theta \, d\theta \\ &= \frac{1}{4}\sin\theta\cos^3\theta + \frac{3}{4}\left(\frac{1}{2}\sin\theta\cos\theta + \frac{1}{2}\int \cos^0\theta \, d\theta\right) \qquad \cos^0\theta = 1! \\ &= \frac{1}{4}\sin\theta\cos^3\theta + \frac{3}{8}\sin\theta\cos\theta + \frac{3}{8}\theta \end{aligned}$$

Next, let's use the following right triangle to get this back to x's.

$$\tan\theta = \frac{u}{\sqrt{2}} = \frac{x-3}{\sqrt{2}} \qquad \sin\theta = \frac{x-3}{\sqrt{(x-3)^2+2}} \qquad \cos\theta = \frac{\sqrt{2}}{\sqrt{(x-3)^2+2}}$$



The cosine integral is then,

$$\begin{aligned} \int \cos^4\theta \, d\theta &= \frac{1}{4}\frac{2\sqrt{2}\,(x-3)}{\left((x-3)^2+2\right)^2} + \frac{3}{8}\frac{\sqrt{2}\,(x-3)}{(x-3)^2+2} + \frac{3}{8}\tan^{-1}\left(\frac{x-3}{\sqrt{2}}\right) \\ &= \frac{\sqrt{2}}{2}\frac{x-3}{\left((x-3)^2+2\right)^2} + \frac{3\sqrt{2}}{8}\frac{x-3}{(x-3)^2+2} + \frac{3}{8}\tan^{-1}\left(\frac{x-3}{\sqrt{2}}\right) \end{aligned}$$

All told then the original integral is,

$$\int \frac{x}{\left(x^2-6x+11\right)^3}\,dx = -\frac{1}{4}\frac{1}{\left(\left(x-3\right)^2+2\right)^2} +$$

$$\frac{3\sqrt{2}}{8}\left(\frac{\sqrt{2}}{2}\frac{x-3}{\left(\left(x-3\right)^2+2\right)^2} + \frac{3\sqrt{2}}{8}\frac{x-3}{\left(x-3\right)^2+2} + \frac{3}{8}\tan^{-1}\left(\frac{x-3}{\sqrt{2}}\right)\right)$$

$$= \frac{1}{8}\frac{3x-11}{\left(\left(x-3\right)^2+2\right)^2} + \frac{9}{32}\frac{x-3}{\left(x-3\right)^2+2} + \frac{9\sqrt{2}}{64}\tan^{-1}\left(\frac{x-3}{\sqrt{2}}\right) + c$$

It's a long and messy answer, but there it is.

Example 5 Evaluate the following integral.

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

Solution

As with the other problems we'll first complete the square on the denominator.

$$4-2x-x^2 = -\left(x^2+2x-4\right) = -\left(x^2+2x+1-1-4\right) = -\left(\left(x+1\right)^2-5\right) = 5-\left(x+1\right)^2$$

The integral is,

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

Now, let's do the substitution.

$$u = x+1 \qquad\qquad x = u-1 \qquad\qquad dx = du$$

and the integral is now,

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

$$= \int \frac{u}{\left(5-u^2\right)^2}\,du - \int \frac{4}{\left(5-u^2\right)^2}\,du$$

In the first integral we'll use the substitution

$$v = 5-u^2$$

and in the second integral we'll use the following trig substitution

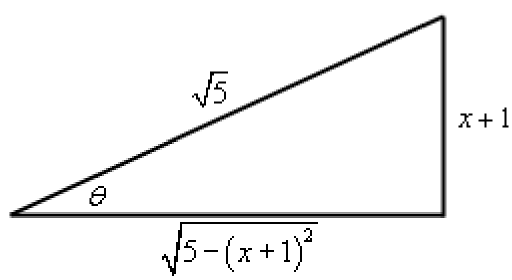
$$u = \sqrt{5}\sin\theta \qquad\qquad du = \sqrt{5}\cos\theta\, d\theta$$

Using these substitutions the integral becomes,

$$\int \frac{x-3}{\left(4-2x-x^2\right)^2}\,dx = -\frac{1}{2}\int \frac{1}{v^2}\,dv - \int \frac{4}{\left(5-5\sin^2\theta\right)^2}\left(\sqrt{5}\cos\theta\right)d\theta$$
$$= \frac{1}{2}\frac{1}{v} - \frac{4\sqrt{5}}{25}\int \frac{\cos\theta}{\left(1-\sin^2\theta\right)^2}\,d\theta$$
$$= \frac{1}{2}\frac{1}{v} - \frac{4\sqrt{5}}{25}\int \frac{\cos\theta}{\cos^4\theta}\,d\theta$$
$$= \frac{1}{2}\frac{1}{v} - \frac{4\sqrt{5}}{25}\int \sec^3\theta\,d\theta$$
$$= \frac{1}{2}\frac{1}{v} - \frac{2\sqrt{5}}{25}\left(\sec\theta\tan\theta + \ln\left|\sec\theta+\tan\theta\right|\right) + c$$

We'll need the following right triangle to finish this integral out.

$$\sin\theta = \frac{u}{\sqrt{5}} = \frac{x+1}{\sqrt{5}} \qquad \sec\theta = \frac{\sqrt{5}}{\sqrt{5-\left(x+1\right)^2}} \qquad \tan\theta = \frac{x+1}{\sqrt{5-\left(x+1\right)^2}}$$



So, going back to x's the integral becomes,

$$\int \frac{x-3}{\left(4-2x-x^2\right)^2}\,dx = \frac{1}{2}\frac{1}{5-u^2} - \frac{2\sqrt{5}}{25}\left(\frac{\sqrt{5}\left(x+1\right)}{5-\left(x+1\right)^2} + \ln\left|\frac{\sqrt{5}}{\sqrt{5-\left(x+1\right)^2}} + \frac{x+1}{\sqrt{5-\left(x+1\right)^2}}\right|\right) + c$$
$$= \frac{1}{10}\frac{4x-1}{5-\left(x+1\right)^2} - \frac{2\sqrt{5}}{25}\ln\left|\frac{x+1+\sqrt{5}}{\sqrt{5-\left(x+1\right)^2}}\right| + c$$

Often the following formula is needed when using the trig substitution that we used in the previous example.

$$\int \sec^m\theta\,d\theta = \frac{1}{m-1}\tan\theta\sec^{m-2}\theta + \frac{m-2}{m-1}\int \sec^{m-2}\theta\,d\theta$$

Note that we'll only need the two trig substitutions that we used here. The third trig substitution that we used will not be needed here.

Integration Strategy

We've now seen a fair number of different integration techniques and so we should probably pause at this point and talk a little bit about a strategy to use for determining the correct technique to use when faced with an integral.

There are a couple of points that need to be made about this strategy. First, it isn't a hard and fast set of rules for determining the method that should be used. It is really nothing more than a general set of guidelines that will help us to identify techniques that may work. Some integrals can be done in more than one way and so depending on the path you take through the strategy you may end up with a different technique than somebody else who also went through this strategy.

Second, while the strategy is presented as a way to identify the technique that could be used on an integral also keep in mind that, for many integrals, it can also automatically exclude certain techniques as well. When going through the strategy keep two lists in mind. The first list is integration techniques that simply won't work and the second list is techniques that look like they might work. After going through the strategy and the second list has only one entry then that is the technique to use. If, on the other hand, there is more than one possible technique to use we will then have to decide on which is liable to be the best for us to use. Unfortunately there is no way to teach which technique is the best as that usually depends upon the person and which technique they find to be the easiest.

Third, don't forget that many integrals can be evaluated in multiple ways and so more than one technique may be used on it. This has already been mentioned in each of the previous points, but is important enough to warrant a separate mention. Sometimes one technique will be significantly easier than the others and so don't just stop at the first technique that appears to work. Always identify all possible techniques and then go back and determine which you feel will be the easiest for you to use.

Next, it's entirely possible that you will need to use more than one method to completely do an integral. For instance a substitution may lead to using integration by parts or partial fractions integral.

Finally, in my class I will accept any valid integration technique as a solution. As already noted there is often more than one way to do an integral and just because I find one technique to be the easiest doesn't mean that you will as well. So, in my class, there is no one right way of doing an integral. You may use any integration technique that I've taught you in this class or you learned in Calculus I to evaluate integrals in this class. In other words, always take the approach that you find to be the easiest.

Note that this final point is more geared towards my class and it's completely possible that your instructor may not agree with this and so be careful in applying this point if you aren't in my class.

Okay, let's get on with the strategy.

1. **Simplify the integrand, if possible.** This step is very important in the integration process. Many integrals can be taken from impossible or very difficult to very easy with a little simplification or manipulation. Don't forget basic trig and algebraic identities as these can often be used to simplify the integral.

   We used this idea when we were looking at integrals involving trig functions. For example consider the following integral.
   $$\int \cos^2 x\,dx$$

   This integral can't be done as is, however simply by recalling the identity,
   $$\cos^2 x = \frac{1}{2}\left(1+\cos\left(2x\right)\right)$$
   the integral becomes very easy to do.

   Note that this example also shows that simplification does not necessarily mean that we'll write the integrand in a "simpler" form. It only means that we'll write the integrand into a form that we can deal with and this is often longer and/or "messier" than the original integral.

2. **See if a "simple" substitution will work.** Look to see if a simple substitution can be used instead of the often more complicated methods from Calculus II. For example consider both of the following integrals.
   $$\int \frac{x}{x^2-1}\,dx \qquad\qquad \int x\sqrt{x^2-1}\,dx$$

   The first integral can be done with partial fractions and the second could be done with a trig substitution.

   However, both could also be evaluated using the substitution $u = x^2 - 1$ and the work involved in the substitution would be significantly less than the work involved in either partial fractions or trig substitution.

   So, always look for quick, simple substitutions before moving on to the more complicated Calculus II techniques.

3. **Identify the type of integral.** Note that any integral may fall into more than one of these types. Because of this fact it's usually best to go all the way through the list and identify all possible types since one may be easier than the other and it's entirely possible that the easier type is listed lower in the list.

   a. Is the integrand a rational expression (*i.e* is the integrand a polynomial divided by a polynomial)? If so, then partial fractions may work on the integral.
   b. Is the integrand a polynomial times a trig function, exponential, or logarithm? If so, then integration by parts may work.
   c. Is the integrand a product of sines and cosines, secant and tangents, or cosecants and cotangents? If so, then the topics from the second section may work. Likewise, don't forget that some quotients involving these functions can also be done using these techniques.

d. Does the integrand involve $\sqrt{b^2x^2+a^2}$, $\sqrt{b^2x^2-a^2}$, or $\sqrt{a^2-b^2x^2}$? If so, then a trig substitution might work nicely.
e. Does the integrand have roots other than those listed above in it? If so, then the substitution $u=\sqrt[n]{g(x)}$ might work.
f. Does the integrand have a quadratic in it? If so, then completing the square on the quadratic might put it into a form that we can deal with.

4. **Can we relate the integral to an integral we already know how to do?** In other words, can we use a substitution or manipulation to write the integrand into a form that does fit into the forms we've looked at previously in this chapter.

   A typical example here is the following integral.
$$\int \cos x\sqrt{1+\sin^2 x}\,dx$$

   This integral doesn't obviously fit into any of the forms we looked at in this chapter. However, with the substitution $u=\sin x$ we can reduce the integral to the form,
$$\int \sqrt{1+u^2}\,du$$
   which is a trig substitution problem.

5. **Do we need to use multiple techniques?** In this step we need to ask ourselves if it is possible that we'll need to use multiple techniques. The example in the previous part is a good example. Using a substitution didn't allow us to actually do the integral. All it did was put the integral and put it into a form that we could use a different technique on.

   Don't ever get locked into the idea that an integral will only require one step to completely evaluate it. Many will require more than one step.

6. **Try again.** If everything that you've tried to this point doesn't work then go back through the process and try again. This time try a technique that that you didn't use the first time around.

As noted above this strategy is not a hard and fast set of rules. It is only intended to guide you through the process of best determining how to do any given integral. Note as well that the only place Calculus II actually arises is in the third step. Steps 1, 2 and 4 involve nothing more than manipulation of the integrand either through direct manipulation of the integrand or by using a substitution. The last two steps are simply ideas to think about in going through this strategy.

Many students go through this process and concentrate almost exclusively on Step 3 (after all this is Calculus II, so it's easy to see why they might do that….) to the exclusion of the other steps. One very large consequence of that exclusion is that often a simple manipulation or substitution is overlooked that could make the integral very easy to do.

Before moving on to the next section we should work a couple of quick problems illustrating a couple of not so obvious simplifications/manipulations and a not so obvious substitution.

Example 1 Evaluate the following integral.

$$\int \frac{\tan x}{\sec^4 x}\,dx$$

Solution

This integral almost falls into the form given in **3c**. It is a quotient of tangent and secant and we know that sometimes we can use the same methods for products of tangents and secants on quotients.

The process from that section tells us that if we have even powers of secant to strip two of them off and convert the rest to tangents. That won't work here. We can split two secants off, but they would be in the denominator and they won't do us any good there. Remember that the point of splitting them off is so they would be there for the substitution $u = \tan x$. That requires them to be in the numerator. So, that won't work and so we'll have to find another solution method.

There are in fact two solution methods to this integral depending on how you want to go about it. We'll take a look at both.

Solution 1

In this solution method we could just convert everything to sines and cosines and see if that gives us an integral we can deal with.

$$\begin{aligned} \int \frac{\tan x}{\sec^4 x}\,dx &= \int \frac{\sin x}{\cos x}\cos^4 x\,dx \\ &= \int \sin x\cos^3 x\,dx \qquad\qquad u = \cos x \\ &= -\int u^3\,du \\ &= -\frac{1}{4}\cos^4 x + c \end{aligned}$$

Note that just converting to sines and cosines won't always work and if it does it won't always work this nicely. Often there will be a lot more work that would need to be done to complete the integral.

Solution 2

This solution method goes back to dealing with secants and tangents. Let's notice that if we had a secant in the numerator we could just use $u = \sec x$ as a substitution and it would be a fairly quick and simple substitution to use. We don't have a secant in the numerator. However we could very easily get a secant in the numerator simply by multiplying the numerator and denominator by secant.

$$\begin{aligned} \int \frac{\tan x}{\sec^4 x}\,dx &= \int \frac{\tan x\sec x}{\sec^5 x}\,dx \qquad\qquad u = \sec x \\ &= \int \frac{1}{u^5}\,du \\ &= -\frac{1}{4}\frac{1}{\sec^4 x} + c \\ &= -\frac{1}{4}\cos^4 x + c \end{aligned}$$

In the previous example we saw two "simplifications" that allowed us to do the integral. The first was using identities to rewrite the integral into terms we could deal with and the second involved multiplying the numerator and the denominator by something to again put the integral into terms we could deal with.

Using identities to rewrite an integral is an important "simplification" and we should not forget about it. Integrals can often be greatly simplified or at least put into a form that can be dealt with by using an identity.

The second "simplification" is not used as often, but does show up on occasion so again, it's best to not forget about it. In fact, let's take another look at an example in which multiplying the numerator and denominator by something will allow us to do an integral.

Example 2 Evaluate the following integral.

$$\int \frac{1}{1+\sin x}\,dx$$

Solution

This is an integral in which if we just concentrate on the third step we won't get anywhere. This integral doesn't appear to be any of the kinds of integrals that we worked in this chapter.

We can do the integral however, if we do the following,

$$\begin{aligned}\int \frac{1}{1+\sin x}\,dx &= \int \frac{1}{1+\sin x}\frac{1-\sin x}{1-\sin x}\,dx \\ &= \int \frac{1-\sin x}{1-\sin^2 x}\,dx\end{aligned}$$

This does not appear to have done anything for us. However, if we now remember the first "simplification" we looked at above we will notice that we can use an identity to rewrite the denominator. Once we do that we can further reduce the integral into something we can deal with.

$$\begin{aligned}\int \frac{1}{1+\sin x}\,dx &= \int \frac{1-\sin x}{\cos^2 x}\,dx \\ &= \int \frac{1}{\cos^2 x} - \frac{\sin x}{\cos x}\frac{1}{\cos x}\,dx \\ &= \int \sec^2 x - \tan x \sec x\,dx \\ &= \tan x - \sec x + c\end{aligned}$$

So, we've seen once again that multiplying the numerator and denominator by something can put the integral into a form that we can integrate. Notice as well that this example also showed that "simplifications" do not necessarily put an integral into a simpler form. They only put the integral into a form that is easier to integrate.

Let's now take a quick look at an example of a substitution that is not so obvious.

Example 3 Evaluate the following integral.

$$\int \cos\left(\sqrt{x}\right) dx$$

Solution

We introduced this example saying that the substitution was not so obvious. However, this is really an integral that falls into the form given by 3e in our strategy above. However, many people miss that form and so don't think about it. So, let's try the following substitution.

$$u = \sqrt{x} \qquad x = u^2 \qquad dx = 2u\,du$$

With this substitution the integral becomes,

$$\int \cos\left(\sqrt{x}\right) dx = 2\int u \cos u \, du$$

This is now an integration by parts integral. Remember that often we will need to use more than one technique to completely do the integral. This is a fairly simple integration by parts problem so I'll leave the remainder of the details to you to check.

$$\int \cos\left(\sqrt{x}\right) dx = 2\left(\cos\left(\sqrt{x}\right) + \sqrt{x}\sin\left(\sqrt{x}\right)\right) + c$$

Before leaving this section we should also point out that there are integrals out there in the world that just can't be done in terms of functions that we know. Some examples of these are.

$$\int \mathbf{e}^{-x^2}\, dx \qquad \int \cos\left(x^2\right) dx \qquad \int \frac{\sin\left(x\right)}{x}\, dx \qquad \int \cos\left(\mathbf{e}^x\right) dx$$

That doesn't mean that these integrals can't be done at some level. If you go to a computer algebra system such as Maple or Mathematica and have it do these integrals it will return the following.

$$\int \mathbf{e}^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}\operatorname{erf}\left(x\right)$$

$$\int \cos\left(x^2\right) dx = \sqrt{\frac{\pi}{2}}\,\operatorname{FresnelC}\left(x\sqrt{\frac{2}{\pi}}\right)$$

$$\int \frac{\sin\left(x\right)}{x}\, dx = \operatorname{Si}\left(x\right)$$

$$\int \cos\left(\mathbf{e}^x\right) dx = \operatorname{Ci}\left(\mathbf{e}^x\right)$$

So it appears that these integrals can in fact be done. However this is a little misleading. Here are the definitions of each of the functions given above.

Error Function

$$\operatorname{erf}\left(x\right) = \frac{2}{\sqrt{\pi}}\int_0^x \mathbf{e}^{-t^2}\, dt$$

The Sine Integral

$$\operatorname{Si}\left(x\right) = \int_0^x \frac{\sin t}{t}\, dt$$

The Fresnel Cosine Integral

$$\text{FresnelC}(x) = \int_0^x \cos\left(\frac{\pi}{2}t^2\right)dt$$

The Cosine Integral

$$\text{Ci}(x) = \gamma + \ln(x) + \int_0^x \frac{\cos t - 1}{t}\,dt$$

Where γ is the Euler-Mascheroni constant.

Note that the first three are simply defined in terms of themselves and so when we say we can integrate them all we are really doing is renaming the integral. The fourth one is a little different and yet it is still defined in terms of an integral that can't be done in practice.

It will be possible to integrate every integral given in this class, but it is important to note that there are integrals that just can't be done. We should also note that after we look at Series we will be able to write down series representations of each of the integrals above.

Improper Integrals

In this section we need to take a look at a couple of different kinds of integrals. Both of these are examples of integrals that are called Improper Integrals.

Let's start with the first kind of improper integrals that we're going to take a look at.

Infinite Interval

In this kind of integral one or both of the limits of integration are infinity. In these cases the interval of integration is said to be over an infinite interval.

Let's take a look at an example that will also show us how we are going to deal with these integrals.

Example 1 Evaluate the following integral.

$$\int_1^\infty \frac{1}{x^2}\,dx$$

Solution

This is an innocent enough looking integral. However, because infinity is not a real number we can't just integrate as normal and then "plug in" the infinity to get an answer.

To see how we're going to do this integral let's think of this as an area problem. So instead of asking what the integral is, let's instead ask what the area under $f(x) = \frac{1}{x^2}$ on the interval $[1,\infty)$ is.

We still aren't able to do this, however, let's step back a little and instead ask what the area under $f(x)$ is on the interval $[1,t]$ where $t > 1$ and *t* is finite. This is a problem that we can do.

$$A_t = \int_1^t \frac{1}{x^2}\,dx = -\left.\frac{1}{x}\right|_1^t = 1 - \frac{1}{t}$$

Now, we can get the area under $f(x)$ on $[1,\infty)$ simply by taking the limit of A_t as *t* goes to infinity.

$$A = \lim_{t\to\infty} A_t = \lim_{t\to\infty}\left(1 - \frac{1}{t}\right) = 1$$

This is then how we will do the integral itself.

$$\begin{aligned}\int_1^\infty \frac{1}{x^2}\,dx &= \lim_{t\to\infty}\int_1^t \frac{1}{x^2}\,dx \\ &= \lim_{t\to\infty}\left.\left(-\frac{1}{x}\right)\right|_1^t \\ &= \lim_{t\to\infty}\left(1 - \frac{1}{t}\right) = 1\end{aligned}$$

So, this is how we will deal with these kinds of integrals in general. We will replace the infinity with a variable (usually *t*), do the integral and then take the limit of the result as *t* goes to infinity.

On a side note, notice that the area under a curve on an infinite interval was not infinity as we might have suspected it to be. In fact, it was a surprisingly small number. Of course this won't always be the case, but it is important enough to point out that not all areas on an infinite interval will yield infinite areas.

Let's now get some definitions out of the way. We will call these integrals **convergent** if the associated limit exists and is a finite number (*i.e.* it's not plus or minus infinity) and **divergent** if the associated limit either doesn't exist or is (plus or minus) infinity.

Let's now formalize up the method for dealing with infinite intervals. There are essentially three cases that we'll need to look at.

1. If $\int_a^t f(x)\,dx$ exists for every $t > a$ then,

$$\int_a^{\infty} f(x)\,dx = \lim_{t\to\infty}\int_a^t f(x)\,dx$$

provided the limit exists and is finite.

2. If $\int_t^b f(x)\,dx$ exists for every $t < b$ then,

$$\int_{-\infty}^{b} f(x)\,dx = \lim_{t\to-\infty}\int_t^b f(x)\,dx$$

provided the limits exists and is finite.

3. If $\int_{-\infty}^{c} f(x)\,dx$ and $\int_c^{\infty} f(x)\,dx$ are both convergent then,

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_c^{\infty} f(x)\,dx$$

Where *c* is any number. Note as well that this requires BOTH of the integrals to be convergent in order for this integral to also be convergent. If either of the two integrals is divergent then so is this integral.

Let's take a look at a couple more examples.

Example 2 Determine if the following integral is convergent or divergent and if it's convergent find its value.

$$\int_1^{\infty} \frac{1}{x}\,dx$$

Solution

So, the first thing we do is convert the integral to a limit.

$$\int_1^{\infty} \frac{1}{x}\,dx = \lim_{t\to\infty}\int_1^t \frac{1}{x}\,dx$$

Now, do the integral and the limit.

$$\int_1^\infty \frac{1}{x}\,dx = \lim_{t\to\infty} \ln(x)\Big|_1^t$$
$$= \lim_{t\to\infty}\left(\ln(t) - \ln 1\right)$$
$$= \infty$$

So, the limit is infinite and so the integral is divergent.

If we go back to thinking in terms of area notice that the area under $g(x) = \frac{1}{x}$ on the interval $[1,\infty)$ is infinite. This is in contrast to the area under $f(x) = \frac{1}{x^2}$ which was quite small. There really isn't all that much difference between these two functions and yet there is a large difference in the area under them. We can actually extend this out to the following fact.

Fact

If $a > 0$ then

$$\int_a^\infty \frac{1}{x^p}\,dx$$

is convergent if $p > 1$ and divergent if $p \le 1$.

One thing to note about this fact is that it's in essence saying that if an integrand goes to zero fast enough then the integral will converge. How fast is fast enough? If we use this fact as a guide it looks like integrands that go to zero faster than $\frac{1}{x}$ goes to zero will probably converge.

Let's take a look at a couple more examples.

Example 3 Determine if the following integral is convergent or divergent. If it is convergent find its value.

$$\int_{-\infty}^0 \frac{1}{\sqrt{3-x}}\,dx$$

Solution

There really isn't much to do with these problems once you know how to do them. We'll convert the integral to a limit/integral pair, evaluate the integral and then the limit.

$$\int_{-\infty}^0 \frac{1}{\sqrt{3-x}}\,dx = \lim_{t\to-\infty}\int_t^0 \frac{1}{\sqrt{3-x}}\,dx$$
$$= \lim_{t\to-\infty} -2\sqrt{3-x}\Big|_t^0$$
$$= \lim_{t\to-\infty}\left(-2\sqrt{3} + 2\sqrt{3-t}\right)$$
$$= -2\sqrt{3} + \infty$$
$$= \infty$$

So, the limit is infinite and so this integral is divergent.

Example 4 Determine if the following integral is convergent or divergent. If it is convergent find its value.

$$\int_{-\infty}^{\infty} x\mathbf{e}^{-x^2}\, dx$$

Solution

In this case we've got infinities in both limits and so we'll need to split the integral up into two separate integrals. We can split the integral up at any point, so let's choose $a = 0$ since this will be a convenient point for the evaluation process. The integral is then,

$$\int_{-\infty}^{\infty} x\mathbf{e}^{-x^2}\, dx = \int_{-\infty}^{0} x\mathbf{e}^{-x^2}\, dx + \int_{0}^{\infty} x\mathbf{e}^{-x^2}\, dx$$

We've now got to look at each of the individual limits.

$$\begin{aligned}\int_{-\infty}^{0} x\mathbf{e}^{-x^2}\, dx &= \lim_{t\to-\infty}\int_{t}^{0} x\mathbf{e}^{-x^2}\, dx\\ &= \lim_{t\to-\infty}\left(-\frac{1}{2}\mathbf{e}^{-x^2}\right)\Bigg|_{t}^{0}\\ &= \lim_{t\to-\infty}\left(-\frac{1}{2}+\frac{1}{2}\mathbf{e}^{-t^2}\right)\\ &= -\frac{1}{2}\end{aligned}$$

So, the first integral is convergent. Note that this does NOT mean that the second integral will also be convergent. So, let's take a look at that one.

$$\begin{aligned}\int_{0}^{\infty} x\mathbf{e}^{-x^2}\, dx &= \lim_{t\to\infty}\int_{0}^{t} x\mathbf{e}^{-x^2}\, dx\\ &= \lim_{t\to\infty}\left(-\frac{1}{2}\mathbf{e}^{-x^2}\right)\Bigg|_{0}^{t}\\ &= \lim_{t\to\infty}\left(-\frac{1}{2}\mathbf{e}^{-t^2}+\frac{1}{2}\right)\\ &= \frac{1}{2}\end{aligned}$$

This integral is convergent and so since they are both convergent the integral we were actually asked to deal with is also convergent and its value is,

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

Example 5 Determine if the following integral is convergent or divergent. If it is convergent find its value.

$$\int_{-2}^{\infty} \sin x\, dx$$

Solution

First convert to a limit.

$$\begin{aligned}\int_{-2}^{\infty} \sin x\, dx &= \lim_{t\to\infty} \int_{-2}^{t} \sin x\, dx \\ &= \lim_{t\to\infty} \left(-\cos x\right)\Big|_{-2}^{t} \\ &= \lim_{t\to\infty} \left(\cos 2 - \cos t\right)\end{aligned}$$

This limit doesn't exist and so the integral is divergent.

In most examples in a Calculus II class that are worked over infinite intervals the limit either exists or is infinite. However, there are limits that don't exist, as the previous example showed, so don't forget about those.

Discontinuous Integrand

We now need to look at the second type of improper integrals that we'll be looking at in this section. These are integrals that have discontinuous integrands. The process here is basically the same with one subtle difference. Here are the general cases that we'll look at for these integrals.

1. If $f(x)$ is continuous on the interval $[a,b)$ and not continuous at $x = b$ then,

$$\int_a^b f(x)\,dx = \lim_{t\to b^-} \int_a^t f(x)\,dx$$

provided the limit exists and is finite. Note as well that we do need to use a left hand limit here since the interval of integration is entirely on the left side of the upper limit.

2. If $f(x)$ is continuous on the interval $(a,b]$ and not continuous at $x = a$ then,

$$\int_a^b f(x)\,dx = \lim_{t\to a^+} \int_t^b f(x)\,dx$$

provided the limit exists and is finite. In this case we need to use a right hand limit here since the interval of integration is entirely on the right side of the lower limit.

3. If $f(x)$ is not continuous at $x = c$ where $a < c < b$ and $\int_a^c f(x)\,dx$ and $\int_c^b f(x)\,dx$ are both convergent then,

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

As with the infinite interval case this requires BOTH of the integrals to be convergent in order for this integral to also be convergent. If either of the two integrals is divergent then so is this integral.

4. If $f(x)$ is not continuous at $x = a$ and $x = b$ and if $\int_a^c f(x)\,dx$ and $\int_c^b f(x)\,dx$ are both convergent then,

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

Where *c* is any number. Again, this requires BOTH of the integrals to be convergent in order for this integral to also be convergent.

Note that the limits in these cases really do need to be right or left handed limits. Since we will be working inside the interval of integration we will need to make sure that we stay inside that interval. This means that we'll use one-sided limits to make sure we stay inside the interval.

Let's do a couple of examples of these kinds of integrals.

Example 6 Determine if the following integral is convergent or divergent. If it is convergent find its value.

$$\int_0^3 \frac{1}{\sqrt{3-x}}\,dx$$

Solution

The problem point is the upper limit so we are in the first case above.

$$\begin{aligned}\int_0^3 \frac{1}{\sqrt{3-x}}\,dx &= \lim_{t\to 3^-}\int_0^t \frac{1}{\sqrt{3-x}}\,dx\\ &= \lim_{t\to 3^-}\left(-2\sqrt{3-x}\right)\Big|_0^t\\ &= \lim_{t\to 3^-}\left(2\sqrt{3}-2\sqrt{3-t}\right)\\ &= 2\sqrt{3}\end{aligned}$$

The limit exists and is finite and so the integral converges and the integral's value is $2\sqrt{3}$.

Example 7 Determine if the following integral is convergent or divergent. If it is convergent find its value.

$$\int_{-2}^3 \frac{1}{x^3}\,dx$$

Solution

This integrand is not continuous at $x=0$ and so we'll need to split the integral up at that point.

$$\int_{-2}^3 \frac{1}{x^3}\,dx = \int_{-2}^0 \frac{1}{x^3}\,dx + \int_0^3 \frac{1}{x^3}\,dx$$

Now we need to look at each of these integrals and see if they are convergent.

$$\begin{aligned}\int_{-2}^0 \frac{1}{x^3}\,dx &= \lim_{t\to 0^-}\int_{-2}^t \frac{1}{x^3}\,dx\\ &= \lim_{t\to 0^-}\left(-\frac{1}{2x^2}\right)\Bigg|_{-2}^t\\ &= \lim_{t\to 0^-}\left(-\frac{1}{2t^2}+\frac{1}{8}\right)\\ &= -\infty\end{aligned}$$

At this point we're done. One of the integrals is divergent that means the integral that we were asked to look at is divergent. We don't even need to bother with the second integral.

Before leaving this section let's note that we can also have integrals that involve both of these cases. Consider the following integral.

Example 8 Determine if the following integral is convergent or divergent. If it is convergent find its value.

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

Solution

This is an integral over an infinite interval that also contains a discontinuous integrand. To do this integral we'll need to split it up into two integrals. We can split it up anywhere, but pick a value that will be convenient for evaluation purposes.

$$\int_0^\infty \frac{1}{x^2}\,dx = \int_0^1 \frac{1}{x^2}\,dx + \int_1^\infty \frac{1}{x^2}\,dx$$

In order for the integral in the example to be convergent we will need BOTH of these to be convergent. If one or both are divergent then the whole integral will also be divergent.

We know that the second integral is convergent by the fact given in the infinite interval portion above. So, all we need to do is check the first integral.

$$\begin{aligned}\int_0^1 \frac{1}{x^2}\,dx &= \lim_{t\to 0^+}\int_t^1 \frac{1}{x^2}\,dx \\ &= \lim_{t\to 0^+}\left(-\frac{1}{x}\right)\Bigg|_t^1 \\ &= \lim_{t\to 0^+}\left(-1+\frac{1}{t}\right) \\ &= \infty\end{aligned}$$

So, the first integral is divergent and so the whole integral is divergent.

Comparison Test for Improper Integrals

Now that we've seen how to actually compute improper integrals we need to address one more topic about them. Often we aren't concerned with the actual value of these integrals. Instead we might only be interested in whether the integral is convergent or divergent. Also, there will be some integrals that we simply won't be able to integrate and yet we would still like to know if they converge or diverge.

To deal with this we've got a test for convergence or divergence that we can use to help us answer the question of convergence for an improper integral.

We will give this test only for a sub-case of the infinite interval integral, however versions of the test exist for the other sub-cases of the infinite interval integrals as well as integrals with discontinuous integrands.

Comparison Test

If $f(x) \geq g(x) \geq 0$ on the interval $[a,\infty)$ then,

1. If $\int_a^\infty f(x)\,dx$ converges then so does $\int_a^\infty g(x)\,dx$.
2. If $\int_a^\infty g(x)\,dx$ diverges then so does $\int_a^\infty f(x)\,dx$.

Note that if you think in terms of area the Comparison Test makes a lot of sense. If $f(x)$ is larger than $g(x)$ then the area under $f(x)$ must also be larger than the area under $g(x)$.

So, if the area under the larger function is finite (*i.e.* $\int_a^\infty f(x)\,dx$ converges) then the area under the smaller function must also be finite (*i.e.* $\int_a^\infty g(x)\,dx$ converges). Likewise, if the area under the smaller function is infinite (*i.e.* $\int_a^\infty g(x)\,dx$ diverges) then the area under the larger function must also be infinite (*i.e.* $\int_a^\infty f(x)\,dx$ diverges).

Be careful not to misuse this test. If the smaller function converges there is no reason to believe that the larger will also converge (after all infinity is larger than a finite number…) and if the larger function diverges there is no reason to believe that the smaller function will also diverge.

Let's work a couple of examples using the comparison test. Note that all we'll be able to do is determine the convergence of the integral. We won't be able to determine the value of the integrals and so won't even bother with that.

Example 1 Determine if the following integral is convergent or divergent.

$$\int_2^\infty \frac{\cos^2 x}{x^2}\,dx$$

Solution

Let's take a second and think about how the Comparison Test works. If this integral is convergent then we'll need to find a larger function that also converges on the same interval. Likewise, if this integral is divergent then we'll need to find a smaller function that also diverges.

So, it seems like it would be nice to have some idea as to whether the integral converges or diverges ahead of time so we will know whether we will need to look for a larger (and convergent) function or a smaller (and divergent) function.

To get the guess for this function let's notice that the numerator is nice and bounded and simply won't get too large. Therefore, it seems likely that the denominator will determine the convergence/divergence of this integral and we know that

$$\int_2^\infty \frac{1}{x^2}\,dx$$

converges since $p = 2 > 1$ by the fact in the previous section. So let's guess that this integral will converge.

So we now know that we need to find a function that is larger than

$$\frac{\cos^2 x}{x^2}$$

and also converges. Making a fraction larger is actually a fairly simple process. We can either make the numerator larger or we can make the denominator smaller. In this case we can't do a lot about the denominator. However we can use the fact that $0 \le \cos^2 x \le 1$ to make the numerator larger (*i.e.* we'll replace the cosine with something we know to be larger, namely 1). So,

$$\frac{\cos^2 x}{x^2} \le \frac{1}{x^2}$$

Now, as we've already noted

$$\int_2^\infty \frac{1}{x^2}\,dx$$

converges and so by the Comparison Test we know that

$$\int_2^\infty \frac{\cos^2 x}{x^2}\,dx$$

must also converge.

Example 2 Determine if the following integral is convergent or divergent.

$$\int_3^\infty \frac{1}{x + \mathbf{e}^x}\,dx$$

Solution

Let's first take a guess about the convergence of this integral. As noted after the fact in the last section about

$$\int_a^{\infty} \frac{1}{x^p}\,dx$$

if the integrand goes to zero faster than $\frac{1}{x}$ then the integral will probably converge. Now, we've got an exponential in the denominator which is approaching infinity much faster than the *x* and so it looks like this integral should probably converge.

So, we need a larger function that will also converge. In this case we can't really make the numerator larger and so we'll need to make the denominator smaller in order to make the function larger as a whole. We will need to be careful however. There are two ways to do this and only one, in this case only one, of them will work for us.

First, notice that since the lower limit of integration is 3 we can say that $x \geq 3 > 0$ and we know that exponentials are always positive. So, the denominator is the sum of two positive terms and if we were to drop one of them the denominator would get smaller. This would in turn make the function larger.

The question then is which one to drop? Let's first drop the exponential. Doing this gives,

$$\frac{1}{x+\mathbf{e}^x} < \frac{1}{x}$$

This is a problem however, since

$$\int_3^{\infty} \frac{1}{x}\,dx$$

diverges by the fact. We've got a larger function that is divergent. This doesn't say anything about the smaller function. Therefore, we chose the wrong one to drop.

Let's try it again and this time let's drop the *x*.

$$\frac{1}{x+\mathbf{e}^x} < \frac{1}{\mathbf{e}^x} = \mathbf{e}^{-x}$$

Also,

$$\begin{aligned}\int_3^{\infty} \mathbf{e}^{-x}\,dx &= \lim_{t\to\infty}\int_3^{t} \mathbf{e}^{-x}\,dx \\ &= \lim_{t\to\infty}\left(-\mathbf{e}^{-t}+\mathbf{e}^{-3}\right) \\ &= \mathbf{e}^{-3}\end{aligned}$$

So, $\int_3^{\infty} \mathbf{e}^{-x}\,dx$ is convergent. Therefore, by the Comparison test

$$\int_3^{\infty} \frac{1}{x+\mathbf{e}^x}\,dx$$

is also convergent.

Example 3 Determine if the following integral is convergent or divergent.

$$\int_3^\infty \frac{1}{x - \mathbf{e}^{-x}}\,dx$$

Solution

This is very similar to the previous example with a couple of very important differences. First, notice that the exponential now goes to zero as x increases instead of growing larger as it did in the previous example (because of the negative in the exponent). Also note that the exponential is now subtracted off the x instead of added onto it.

The fact that the exponential goes to zero means that this time the x in the denominator will probably dominate the term and that means that the integral probably diverges. We will therefore need to find a smaller function that also diverges.

Making fractions smaller is pretty much the same as making fractions larger. In this case we'll need to either make the numerator smaller or the denominator larger.

This is where the second change will come into play. As before we know that both x and the exponential are positive. However, this time since we are subtracting the exponential from the x if we were to drop the exponential the denominator will become larger and so the fraction will become smaller. In other words,

$$\frac{1}{x - \mathbf{e}^{-x}} > \frac{1}{x}$$

and we know that

$$\int_3^\infty \frac{1}{x}\,dx$$

diverges and so by the Comparison Test we know that

$$\int_3^\infty \frac{1}{x - \mathbf{e}^{-x}}\,dx$$

must also diverge.

Example 4 Determine if the following integral is convergent or divergent.

$$\int_1^\infty \frac{1 + 3\sin^4(2x)}{\sqrt{x}}\,dx$$

Solution

First notice that as with the first example, the numerator in this function is going to be bounded since the sine is never larger than 1. Therefore, since the exponent on the denominator is less than 1 we can guess that the integral will probably diverge. We will need a smaller function that also diverges.

We know that $0 \le \sin^4(2x) \le 1$. In particular, this term is positive and so if we drop it from the numerator the numerator will get smaller. This gives,

$$\frac{1 + 3\sin^4(2x)}{\sqrt{x}} > \frac{1}{\sqrt{x}}$$

and

$$\int_1^\infty \frac{1}{\sqrt{x}}\,dx$$

diverges so by the Comparison Test

$$\int_1^\infty \frac{1+3\sin^4\left(2x\right)}{\sqrt{x}}\,dx$$

also diverges.

Okay, we've seen a few examples of the Comparison Test now. However, most of them worked pretty much the same way. All the functions were rational and all we did for most of them was add or subtract something from the numerator or denominator to get what we want.

Let's take a look at an example that works a little differently so we don't get too locked into these ideas.

Example 5 Determine if the following integral is convergent or divergent.

$$\int_1^\infty \frac{\mathbf{e}^{-x}}{x}\,dx$$

Solution

Normally, the presence of just an *x* in the denominator would lead us to guess divergent for this integral. However, the exponential in the numerator will approach zero so fast that instead we'll need to guess that this integral converges.

To get a larger function we'll use the fact that we know from the limits of integration that $x > 1$. This means that if we just replace the *x* in the denominator with 1 (which is always smaller than *x*) we will make the denominator smaller and so the function will get larger.

$$\frac{\mathbf{e}^{-x}}{x} < \frac{\mathbf{e}^{-x}}{1} = \mathbf{e}^{-x}$$

and we can show that

$$\int_1^\infty \mathbf{e}^{-x}\,dx$$

converges. In fact, we've already done this for a lower limit of 3 and changing that to a 1 won't change the convergence of the integral. Therefore, by the Comparison Test

$$\int_1^\infty \frac{\mathbf{e}^{-x}}{x}\,dx$$

also converges.

We should also really work an example that doesn't involve a rational function since there is no reason to assume that we'll always be working with rational functions.

Example 6 Determine if the following integral is convergent or divergent.

$$\int_1^\infty \mathbf{e}^{-x^2}\,dx$$

Solution

We know that exponentials with negative exponents die down to zero very fast so it makes sense to guess that this integral will be convergent. We need a larger function, but this time we don't have a fraction to work with so we'll need to do something different.

We'll take advantage of the fact that $\mathbf{e}^{-x}$ is a decreasing function. This means that

$$x_1 > x_2 \qquad \Rightarrow \qquad \mathbf{e}^{-x_1} < \mathbf{e}^{-x_2}$$

In other words, plug in a larger number and the function gets smaller.

From the limits of integration we know that $x > 1$ and this means that if we square x it will get larger. Or,

$$x^2 > x \qquad \text{provided } x > 1$$

Note that we can only say this since $x > 1$. This won't be true if $x \le 1$! We can now use the fact that $\mathbf{e}^{-x}$ is a decreasing function to get,

$$\mathbf{e}^{-x^2} < \mathbf{e}^{-x}$$

So, $\mathbf{e}^{-x}$ is a larger function than $\mathbf{e}^{-x^2}$ and we know that

$$\int_1^\infty \mathbf{e}^{-x}\,dx$$

converges so by the Comparison Test we also know that

$$\int_1^\infty \mathbf{e}^{-x^2}\,dx$$

is convergent.

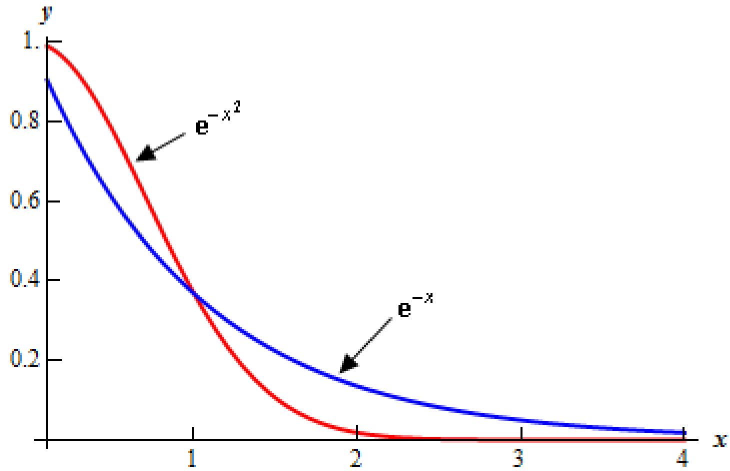
The last two examples made use of the fact that $x > 1$. Let's take a look at an example to see how do we would have to go about these if the lower limit had been smaller than 1.

Example 7 Determine if the following integral is convergent or divergent.

$$\int_{\frac{1}{2}}^\infty \mathbf{e}^{-x^2}\,dx$$

Solution

First, we need to note that $\mathbf{e}^{-x^2} \le \mathbf{e}^{-x}$ is only true on the interval $[1,\infty)$ as is illustrated in the graph below.



So, we can't just proceed as we did in the previous example with the Comparison Test on the interval $\left[\frac{1}{2},\infty\right)$. However, this isn't the problem it might at first appear to be. We can always write the integral as follows,

$$\begin{aligned}\int_{\frac{1}{2}}^{\infty} \mathbf{e}^{-x^2}\,dx &= \int_{\frac{1}{2}}^{1} \mathbf{e}^{-x^2}\,dx + \int_{1}^{\infty} \mathbf{e}^{-x^2}\,dx \\ &= 0.28554 + \int_{1}^{\infty} \mathbf{e}^{-x^2}\,dx\end{aligned}$$

We used Mathematica to get the value of the first integral. Now, if the second integral converges it will have a finite value and so the sum of two finite values will also be finite and so the original integral will converge. Likewise, if the second integral diverges it will either be infinite or not have a value at all and adding a finite number onto this will not all of a sudden make it finite or exist and so the original integral will diverge. Therefore, this integral will converge or diverge depending only on the convergence of the second integral.

As we saw in Example 6 the second integral does converge and so the whole integral must also converge.

As we saw in this example, if we need to, we can split the integral up into one that doesn't involve any problems and can be computed and one that may contain a problem that we can use the Comparison Test on to determine its convergence.

Approximating Definite Integrals

In this chapter we've spent quite a bit of time on computing the values of integrals. However, not all integrals can be computed. A perfect example is the following definite integral.

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

We now need to talk a little bit about estimating values of definite integrals. We will look at three different methods, although one should already be familiar to you from your Calculus I days. We will develop all three methods for estimating

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

by thinking of the integral as an area problem and using known shapes to estimate the area under the curve.

Let's get first develop the methods and then we'll try to estimate the integral shown above.

Midpoint Rule

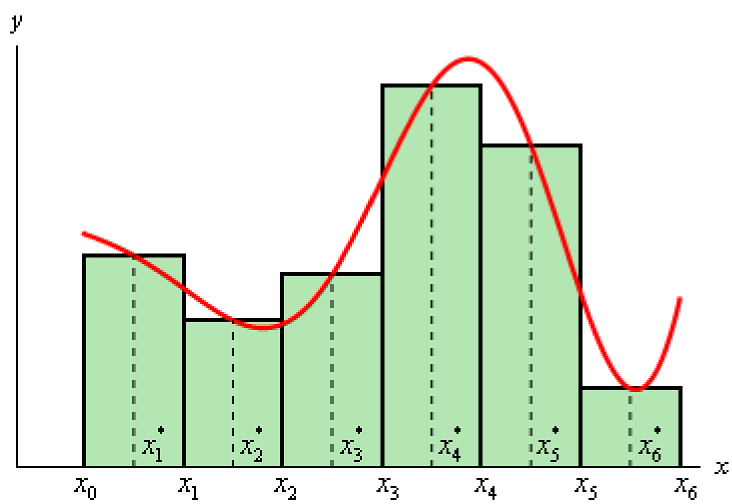
This is the rule that should be somewhat familiar to you. We will divide the interval $[a,b]$ into n subintervals of equal width,

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

We will denote each of the intervals as follows,

$$[x_0, x_1],\ [x_1, x_2], \ldots, [x_{n-1}, x_n] \quad \text{where } x_0 = a \text{ and } x_n = b$$

Then for each interval let x_i^* be the midpoint of the interval. We then sketch in rectangles for each subinterval with a height of $f\left(x_i^*\right)$. Here is a graph showing the set up using $n = 6$.



We can easily find the area for each of these rectangles and so for a general n we get that,

$$\int_a^b f(x)\,dx \approx \Delta x\, f\left(x_1^*\right) + \Delta x\, f\left(x_2^*\right) + \cdots + \Delta x\, f\left(x_n^*\right)$$

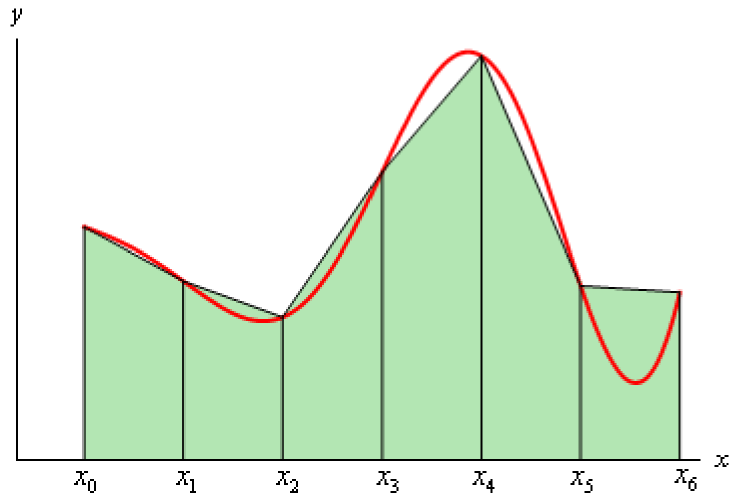
Or, upon factoring out a Δx we get the general Midpoint Rule.

$$\boxed{\int_a^b f(x)\,dx \approx \Delta x\left[f(x_1^*)+f(x_2^*)+\cdots+f(x_n^*)\right]}$$

Trapezoid Rule
For this rule we will do the same set up as for the Midpoint Rule. We will break up the interval $[a,b]$ into *n* subintervals of width,

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

Then on each subinterval we will approximate the function with a straight line that is equal to the function values at either endpoint of the interval. Here is a sketch of this case for $n=6$.



Each of these objects is a trapezoid (hence the rule's name…) and as we can see some of them do a very good job of approximating the actual area under the curve and others don't do such a good job.

The area of the trapezoid in the interval $[x_{i-1},x_i]$ is given by,

$$A_i = \frac{\Delta x}{2}\left(f(x_{i-1})+f(x_i)\right)$$

So, if we use *n* subintervals the integral is approximately,

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

Upon doing a little simplification we arrive at the general Trapezoid Rule.

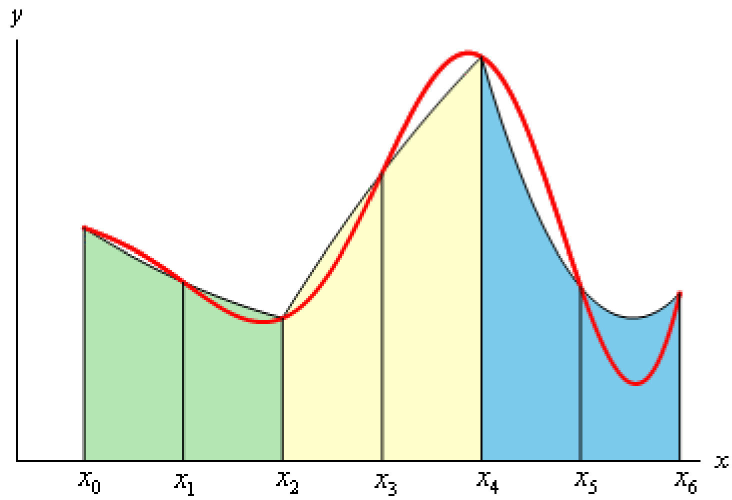
$$\boxed{\int_a^b f(x)\,dx \approx \frac{\Delta x}{2}\left[f(x_0)+2f(x_1)+2f(x_2)+\cdots+2f(x_{n-1})+f(x_n)\right]}$$

Note that all the function evaluations, with the exception of the first and last, are multiplied by 2.

Simpson's Rule
This is the final method we're going to take a look at and in this case we will again divide up the interval $[a,b]$ into n subintervals. However unlike the previous two methods we need to require that n be even. The reason for this will be evident in a bit. The width of each subinterval is,

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

In the Trapezoid Rule we approximated the curve with a straight line. For Simpson's Rule we are going to approximate the function with a quadratic and we're going to require that the quadratic agree with three of the points from our subintervals. Below is a sketch of this using $n=6$. Each of the approximations is colored differently so we can see how they actually work.



Notice that each approximation actually covers two of the subintervals. This is the reason for requiring n to be even. Some of the approximations look more like a line than a quadratic, but they really are quadratics. Also note that some of the approximations do a better job than others. It can be shown that the area under the approximation on the intervals $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$ is,

$$A_i = \frac{\Delta x}{3}\left(f\left(x_{i-1}\right) + 4 f\left(x_i\right) + f\left(x_{i+1}\right)\right)$$

If we use n subintervals the integral is then approximately,

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{3}\left(f\left(x_0\right) + 4 f\left(x_1\right) + f\left(x_2\right)\right) + \frac{\Delta x}{3}\left(f\left(x_2\right) + 4 f\left(x_3\right) + f\left(x_4\right)\right)$$
$$+ \cdots + \frac{\Delta x}{3}\left(f\left(x_{n-2}\right) + 4 f\left(x_{n-1}\right) + f\left(x_n\right)\right)$$

Upon simplifying we arrive at the general Simpson's Rule.

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

In this case notice that all the function evaluations at points with odd subscripts are multiplied by 4 and all the function evaluations at points with even subscripts (except for the first and last) are multiplied by 2. If you can remember this, this is a fairly easy rule to remember.

Okay, it's time to work an example and see how these rules work.

Example 1 Using $n = 4$ and all three rules to approximate the value of the following integral.

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

Solution

First, for reference purposes, Maple gives the following value for this integral.

$$\int_0^2 \mathbf{e}^{x^2}\, dx = 16.45262776$$

In each case the width of the subintervals will be,

$$\Delta x = \frac{2-0}{4} = \frac{1}{2}$$

and so the subintervals will be,

$$[0,\ 0.5],\ [0.5,\ 1],\ [1,\ 1.5],\ [1.5,\ 2]$$

Let's go through each of the methods.

Midpoint Rule

$$\int_0^2 \mathbf{e}^{x^2}\, dx \approx \frac{1}{2}\left(\mathbf{e}^{(0.25)^2} + \mathbf{e}^{(0.75)^2} + \mathbf{e}^{(1.25)^2} + \mathbf{e}^{(1.75)^2}\right) = 14.48561253$$

Remember that we evaluate at the midpoints of each of the subintervals here! The Midpoint Rule has an error of 1.96701523.

Trapezoid Rule

$$\int_0^2 \mathbf{e}^{x^2}\, dx \approx \frac{1/2}{2}\left(\mathbf{e}^{(0)^2} + 2\mathbf{e}^{(0.5)^2} + 2\mathbf{e}^{(1)^2} + 2\mathbf{e}^{(1.5)^2} + \mathbf{e}^{(2)^2}\right) = 20.64455905$$

The Trapezoid Rule has an error of 4.19193129

Simpson's Rule

$$\int_0^2 \mathbf{e}^{x^2}\, dx \approx \frac{1/2}{3}\left(\mathbf{e}^{(0)^2} + 4\mathbf{e}^{(0.5)^2} + 2\mathbf{e}^{(1)^2} + 4\mathbf{e}^{(1.5)^2} + \mathbf{e}^{(2)^2}\right) = 17.35362645$$

The Simpson's Rule has an error of 0.90099869.

None of the estimations in the previous example are all that good. The best approximation in this case is from the Simpson's Rule and yet it still had an error of almost 1. To get a better estimation we would need to use a larger *n*. So, for completeness sake here are the estimates for some larger value of *n*.

n	Midpoint Approx.	Midpoint Error	Trapezoid Approx.	Trapezoid Error	Simpson's Approx.	Simpson's Error
8	15.9056767	0.5469511	17.5650858	1.1124580	16.5385947	0.0859669
16	16.3118539	0.1407739	16.7353812	0.2827535	16.4588131	0.0061853
32	16.4171709	0.0354568	16.5236176	0.0709898	16.4530297	0.0004019
64	16.4437469	0.0088809	16.4703942	0.0177665	16.4526531	0.0000254
128	16.4504065	0.0022212	16.4570706	0.0044428	16.4526294	0.0000016

In this case we were able to determine the error for each estimate because we could get our hands on the exact value. Often this won't be the case and so we'd next like to look at error bounds for each estimate.

These bounds will give the largest possible error in the estimate, but it should also be pointed out that the actual error may be significantly smaller than the bound. The bound is only there so we can say that we know the actual error will be less than the bound.

So, suppose that $\left|f''(x)\right| \le K$ and $\left|f^{(4)}(x)\right| \le M$ for $a \le x \le b$ then if E_M, E_T, and E_S are the actual errors for the Midpoint, Trapezoid and Simpson's Rule we have the following bounds,

$$\left|E_M\right| \le \frac{K(b-a)^3}{24n^2} \qquad \left|E_T\right| \le \frac{K(b-a)^3}{12n^2} \qquad \left|E_S\right| \le \frac{M(b-a)^5}{180n^4}$$

Example 2 Determine the error bounds for the estimations in the last example.

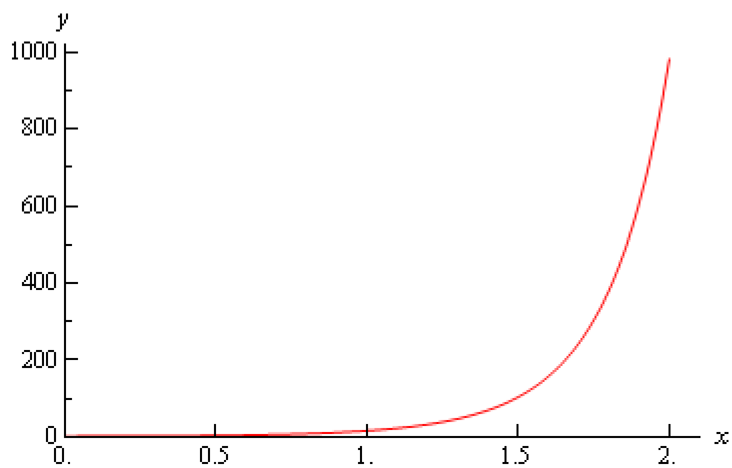
Solution

We already know that $n=4$, $a=0$, and $b=2$ so we just need to compute K (the largest value of the second derivative) and M (the largest value of the fourth derivative). This means that we'll need the second and fourth derivative of $f(x)$.

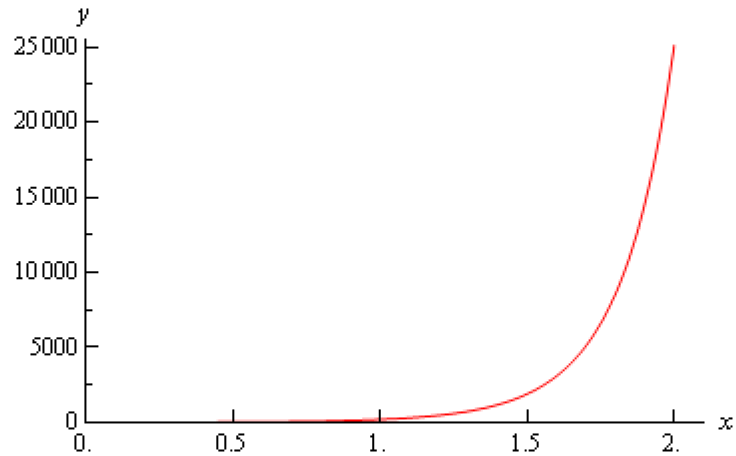
$$f''(x) = 2\mathbf{e}^{x^2}\left(1+2x^2\right)$$

$$f^{(4)}(x) = 4\mathbf{e}^{x^2}\left(3+12x^2+4x^4\right)$$

Here is a graph of the second derivative.



Here is a graph of the fourth derivative.



So, from these graphs it's clear that the largest value of both of these are at $x = 2$. So,

$$f''(2) = 982.76667 \qquad \Rightarrow \qquad K = 983$$

$$f^{(4)}(2) = 25115.14901 \qquad \Rightarrow \qquad M = 25116$$

We rounded to make the computations simpler.

Here are the bounds for each rule.

$$\left|E_M\right| \le \frac{983(2-0)^3}{24(4)^2} = 20.4791666667$$

$$\left|E_T\right| \le \frac{983(2-0)^3}{12(4)^2} = 40.9583333333$$

$$\left|E_S\right| \le \frac{25116(2-0)^5}{180(4)^4} = 17.4416666667$$

In each case we can see that the errors are significantly smaller than the actual bounds.

Applications of Integrals

Introduction

In this section we're going to take a look at some of the applications of integration. It should be noted as well that these applications are presented here, as opposed to Calculus I, simply because many of the integrals that arise from these applications tend to require techniques that we discussed in the previous chapter.

Here is a list of applications that we'll be taking a look at in this chapter.

Arc Length – We'll determine the length of a curve in this section.

Surface Area – In this section we'll determine the surface area of a solid of revolution.

Center of Mass – Here we will determine the center of mass or centroid of a thin plate.

Hydrostatic Pressure and Force – We'll determine the hydrostatic pressure and force on a vertical plate submerged in water.

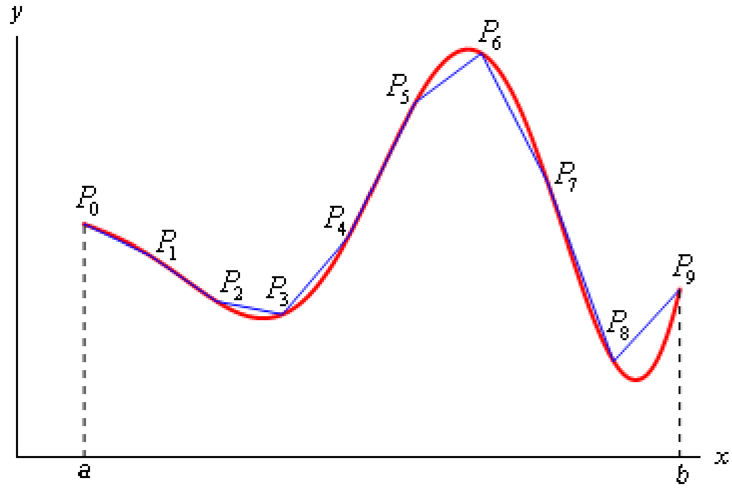
Probability – Here we will look at probability density functions and computing the mean of a probability density function.

Arc Length

In this section we are going to look at computing the arc length of a function. Because it's easy enough to derive the formulas that we'll use in this section we will derive one of them and leave the other to you to derive.

We want to determine the length of the continuous function $y = f(x)$ on the interval $[a,b]$. We'll also need to assume that the derivative is continuous on $[a,b]$.

Initially we'll need to estimate the length of the curve. We'll do this by dividing the interval up into n equal subintervals each of width Δx and we'll denote the point on the curve at each point by P_i. We can then approximate the curve by a series of straight lines connecting the points. Here is a sketch of this situation for $n = 9$.



Now denote the length of each of these line segments by $|P_{i-1}\ P_i|$ and the length of the curve will then be approximately,

$$L \approx \sum_{i=1}^{n} |P_{i-1}\ P_i|$$

and we can get the exact length by taking n larger and larger. In other words, the exact length will be,

$$L = \lim_{n\to\infty} \sum_{i=1}^{n} |P_{i-1}\ P_i|$$

Now, let's get a better grasp on the length of each of these line segments. First, on each segment let's define $\Delta y_i = y_i - y_{i-1} = f(x_i) - f(x_{i-1})$. We can then compute directly the length of the line segments as follows.

$$|P_{i-1}\ P_i| = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} = \sqrt{\Delta x^2 + \Delta y_i^2}$$

By the Mean Value Theorem we know that on the interval $[x_{i-1}, x_i]$ there is a point x_i^* so that,

$$f\left(x_i\right)-f\left(x_{i-1}\right)=f'\left(x_i^*\right)\left(x_i-x_{i-1}\right)$$
$$\Delta y_i = f'\left(x_i^*\right)\Delta x$$

Therefore, the length can now be written as,

$$\begin{aligned}\left|P_{i-1}\ P_i\right| &= \sqrt{\left(x_i-x_{i-1}\right)^2+\left(y_i-y_{i-1}\right)^2}\\ &= \sqrt{\Delta x^2+\left[f'\left(x_i^*\right)\right]^2\Delta x^2}\\ &= \sqrt{1+\left[f'\left(x_i^*\right)\right]^2}\ \ \Delta x\end{aligned}$$

The exact length of the curve is then,

$$\begin{aligned}L &= \lim_{n\to\infty}\sum_{i=1}^{n}\left|P_{i-1}\ P_i\right|\\ &= \lim_{n\to\infty}\sum_{i=1}^{n}\sqrt{1+\left[f'\left(x_i^*\right)\right]^2}\ \ \Delta x\end{aligned}$$

However, using the definition of the definite integral, this is nothing more than,

$$L=\int_a^b\sqrt{1+\left[f'\left(x\right)\right]^2}\,dx$$

A slightly more convenient notation (in my opinion anyway) is the following.

$$L=\int_a^b\sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx$$

In a similar fashion we can also derive a formula for $x=h\left(y\right)$ on $\left[c,d\right]$. This formula is,

$$L=\int_c^d\sqrt{1+\left[h'\left(y\right)\right]^2}\,dy=\int_c^d\sqrt{1+\left(\frac{dx}{dy}\right)^2}\,dy$$

Again, the second form is probably a little more convenient.

Note the difference in the derivative under the square root! Don't get too confused. With one we differentiate with respect to *x* and with the other we differentiate with respect to *y*. One way to keep the two straight is to notice that the differential in the "denominator" of the derivative will match up with the differential in the integral. This is one of the reasons why the second form is a little more convenient.

Before we work any examples we need to make a small change in notation. Instead of having two formulas for the arc length of a function we are going to reduce it, in part, to a single formula.

From this point on we are going to use the following formula for the length of the curve.

Arc Length Formula(s)

$$L = \int ds$$

where,

$$ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \qquad \text{if } y = f\left(x\right),\ a \le x \le b$$

$$ds = \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy \qquad \text{if } x = h\left(y\right),\ c \le y \le d$$

Note that no limits were put on the integral as the limits will depend upon the *ds* that we're using. Using the first *ds* will require *x* limits of integration and using the second *ds* will require *y* limits of integration.

Thinking of the arc length formula as a single integral with different ways to define *ds* will be convenient when we run across arc lengths in future sections. Also, this *ds* notation will be a nice notation for the next section as well.

Now that we've derived the arc length formula let's work some examples.

Example 1 Determine the length of $y = \ln\left(\sec x\right)$ between $0 \le x \le \frac{\pi}{4}$.

Solution

In this case we'll need to use the first *ds* since the function is in the form $y = f\left(x\right)$. So, let's get the derivative out of the way.

$$\frac{dy}{dx} = \frac{\sec x \tan x}{\sec x} = \tan x \qquad\qquad \left(\frac{dy}{dx}\right)^2 = \tan^2 x$$

Let's also get the root out of the way since there is often simplification that can be done and there's no reason to do that inside the integral.

$$\sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \sqrt{1 + \tan^2 x} = \sqrt{\sec^2 x} = \left|\sec x\right| = \sec x$$

Note that we could drop the absolute value bars here since secant is positive in the range given.

The arc length is then,

$$\begin{aligned} L &= \int_0^{\frac{\pi}{4}} \sec x\, dx \\ &= \ln\left|\sec x + \tan x\right|\Big|_0^{\frac{\pi}{4}} \\ &= \ln\left(\sqrt{2} + 1\right) \end{aligned}$$

Example 2 Determine the length of $x = \frac{2}{3}(y-1)^{\frac{3}{2}}$ between $1 \le y \le 4$.

Solution
There is a very common mistake that students make in problems of this type. Many students see that the function is in the form $x = h(y)$ and they immediately decide that it will be too difficult to work with it in that form so they solve for *y* to get the function into the form $y = f(x)$. While that can be done here it will lead to a messier integral for us to deal with.

Sometimes it's just easier to work with functions in the form $x = h(y)$. In fact, if you can work with functions in the form $y = f(x)$ then you can work with functions in the form $x = h(y)$. There really isn't a difference between the two so don't get excited about functions in the form $x = h(y)$.

Let's compute the derivative and the root.

$$\frac{dx}{dy} = (y-1)^{\frac{1}{2}} \qquad \Rightarrow \qquad \sqrt{1+\left(\frac{dx}{dy}\right)^2} = \sqrt{1+y-1} = \sqrt{y}$$

As you can see keeping the function in the form $x = h(y)$ is going to lead to a very easy integral. To see what would happen if we tried to work with the function in the form $y = f(x)$ see the next example.

Let's get the length.

$$\begin{aligned} L &= \int_1^4 \sqrt{y}\,dy \\ &= \frac{2}{3} y^{\frac{3}{2}} \Big|_1^4 \\ &= \frac{14}{3} \end{aligned}$$

As noted in the last example we really do have a choice as to which *ds* we use. Provided we can get the function in the form required for a particular *ds* we can use it. However, as also noted above, there will often be a significant difference in difficulty in the resulting integrals. Let's take a quick look at what would happen in the previous example if we did put the function into the form $y = f(x)$.

Example 3 Redo the previous example using the function in the form $y = f(x)$ instead.

Solution
In this case the function and its derivative would be,

$$y = \left(\frac{3x}{2}\right)^{\frac{2}{3}} + 1 \qquad\qquad \frac{dy}{dx} = \left(\frac{3x}{2}\right)^{-\frac{1}{3}}$$

The root in the arc length formula would then be.

$$\sqrt{1+\left(\frac{dy}{dx}\right)^2} = \sqrt{1+\frac{1}{\left(\frac{3x}{2}\right)^{\frac{2}{3}}}} = \sqrt{\frac{\left(\frac{3x}{2}\right)^{\frac{2}{3}}+1}{\left(\frac{3x}{2}\right)^{\frac{2}{3}}}} = \frac{\sqrt{\left(\frac{3x}{2}\right)^{\frac{2}{3}}+1}}{\left(\frac{3x}{2}\right)^{\frac{1}{3}}}$$

All the simplification work above was just to put the root into a form that will allow us to do the integral.

Now, before we write down the integral we'll also need to determine the limits. This particular *ds* requires *x* limits of integration and we've got *y* limits. They are easy enough to get however. Since we know *x* as a function of *y* all we need to do is plug in the original *y* limits of integration and get the *x* limits of integration. Doing this gives,

$$0 \le x \le \frac{2}{3}(3)^{\frac{3}{2}}$$

Not easy limits to deal with, but there they are.

Let's now write down the integral that will give the length.

$$L = \int_0^{\frac{2}{3}(3)^{\frac{3}{2}}} \frac{\sqrt{\left(\frac{3x}{2}\right)^{\frac{2}{3}}+1}}{\left(\frac{3x}{2}\right)^{\frac{1}{3}}}\,dx$$

That's a really unpleasant looking integral. It can be evaluated however using the following substitution.

$$u = \left(\frac{3x}{2}\right)^{\frac{2}{3}} + 1 \qquad\qquad du = \left(\frac{3x}{2}\right)^{-\frac{1}{3}} dx$$

$$x = 0 \qquad \Rightarrow \qquad u = 1$$

$$x = \frac{2}{3}(3)^{\frac{3}{2}} \qquad \Rightarrow \qquad u = 4$$

Using this substitution the integral becomes,

$$\begin{aligned} L &= \int_1^4 \sqrt{u}\,du \\ &= \left.\frac{2}{3}u^{\frac{3}{2}}\right|_1^4 \\ &= \frac{14}{3} \end{aligned}$$

So, we got the same answer as in the previous example. Although that shouldn't really be all that surprising since we were dealing with the same curve.

From a technical standpoint the integral in the previous example was not that difficult. It was just a Calculus I substitution. However, from a practical standpoint the integral was significantly more difficult than the integral we evaluated in Example 2. So, the moral of the story here is that we can use either formula (provided we can get the function in the correct form of course) however one will often be significantly easier to actually evaluate.

Okay, let's work one more example.

Example 4 Determine the length of $x = \frac{1}{2}y^2$ for $0 \le x \le \frac{1}{2}$. Assume that *y* is positive.

Solution

We'll use the second *ds* for this one as the function is already in the correct form for that one. Also, the other *ds* would again lead to a particularly difficult integral. The derivative and root will then be,

$$\frac{dx}{dy} = y \qquad \Rightarrow \qquad \sqrt{1+\left(\frac{dx}{dy}\right)^2} = \sqrt{1+y^2}$$

Before writing down the length notice that we were given *x* limits and we will need *y* limits for this *ds*. With the assumption that *y* is positive these are easy enough to get. All we need to do is plug *x* into our equation and solve for *y*. Doing this gives,

$$0 \le y \le 1$$

The integral for the arc length is then,

$$L = \int_0^1 \sqrt{1+y^2}\,dy$$

This integral will require the following trig substitution.

$$y = \tan\theta \qquad dy = \sec^2\theta\,d\theta$$

$$y = 0 \quad \Rightarrow \quad 0 = \tan\theta \quad \Rightarrow \quad \theta = 0$$

$$y = 1 \quad \Rightarrow \quad 1 = \tan\theta \quad \Rightarrow \quad \theta = \frac{\pi}{4}$$

$$\sqrt{1+y^2} = \sqrt{1+\tan^2\theta} = \sqrt{\sec^2\theta} = \left|\sec\theta\right| = \sec\theta$$

The length is then,

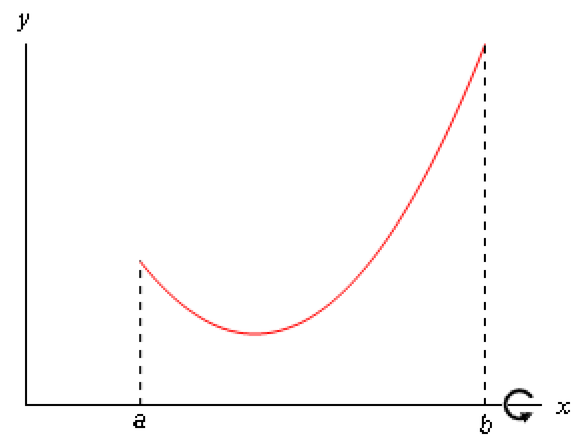
$$\begin{aligned} L &= \int_0^{\frac{\pi}{4}} \sec^3\theta\,d\theta \\ &= \frac{1}{2}\left(\sec\theta\tan\theta + \ln\left|\sec\theta+\tan\theta\right|\right)\Big|_0^{\frac{\pi}{4}} \\ &= \frac{1}{2}\left(\sqrt{2}+\ln\left(1+\sqrt{2}\right)\right) \end{aligned}$$

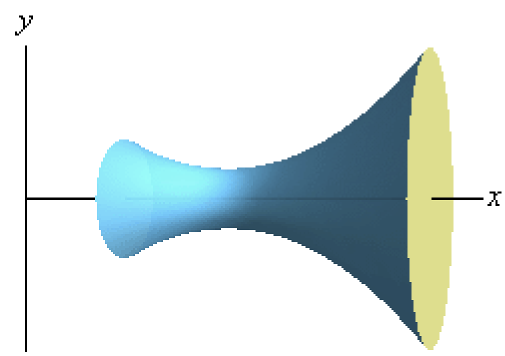
The first couple of examples ended up being fairly simple Calculus I substitutions. However, as this last example had shown we can end up with trig substitutions as well for these integrals.

Surface Area

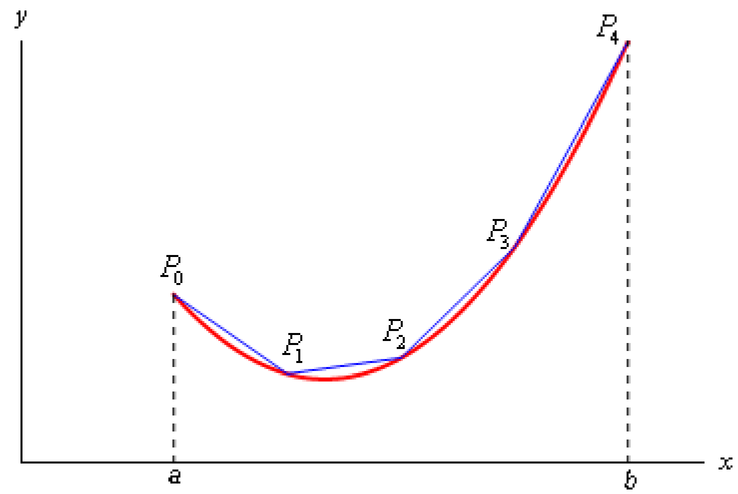
In this section we are going to look once again at solids of revolution. We first looked at them back in Calculus I when we found the volume of the solid of revolution. In this section we want to find the surface area of this region.

So, for the purposes of the derivation of the formula, let's look at rotating the continuous function $y = f\left(x\right)$ in the interval $\left[a,b\right]$ about the *x*-axis. We'll also need to assume that the derivative is continuous on $\left[a,b\right]$. Below is a sketch of a function and the solid of revolution we get by rotating the function about the *x*-axis.

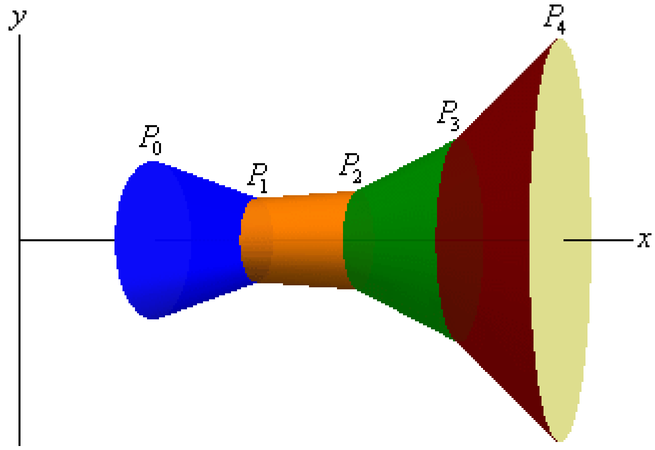




We can derive a formula for the surface area much as we derived the formula for arc length. We'll start by dividing the interval into *n* equal subintervals of width Δx. On each subinterval we will approximate the function with a straight line that agrees with the function at the endpoints of each interval. Here is a sketch of that for our representative function using $n = 4$.



Now, rotate the approximations about the *x*-axis and we get the following solid.



The approximation on each interval gives a distinct portion of the solid and to make this clear each portion is colored differently. Each of these portions are called frustums and we know how to find the surface area of frustums.

The surface area of a frustum is given by,

$$A = 2\pi r l$$

where,

$$r = \frac{1}{2}(r_1 + r_2) \qquad r_1 = \text{radius of right end}$$

$$r_2 = \text{radius of left end}$$

and l is the length of the slant of the frustum.

For the frustum on the interval $[x_{i-1}, x_i]$ we have,

$$r_1 = f(x_i)$$

$$r_2 = f(x_{i-1})$$

$$l = |P_{i-1}\ P_i| \quad \left(\text{length of the line segment connecting } P_i \text{ and } P_{i-1}\right)$$

and we know from the previous section that,

$$|P_{i-1}\ P_i| = \sqrt{1 + \left[f'(x_i^*)\right]^2}\ \Delta x \quad \text{where } x_i^* \text{ is some point in } [x_{i-1}, x_i]$$

Before writing down the formula for the surface area we are going to assume that Δx is "small" and since $f(x)$ is continuous we can then assume that,

$$f(x_i) \approx f(x_i^*) \qquad \text{and} \qquad f(x_{i-1}) \approx f(x_i^*)$$

So, the surface area of the frustum on the interval $[x_{i-1}, x_i]$ is approximately,

$$\begin{aligned} A_i &= 2\pi\left(\frac{f(x_i)+f(x_{i-1})}{2}\right)\left|P_{i-1}\ P_i\right| \\ &\approx 2\pi f(x_i^*)\sqrt{1+\left[f'(x_i^*)\right]^2}\ \ \Delta x \end{aligned}$$

The surface area of the whole solid is then approximately,

$$S \approx \sum_{i=1}^{n} 2\pi f(x_i^*)\sqrt{1+\left[f'(x_i^*)\right]^2}\ \ \Delta x$$

and we can get the exact surface area by taking the limit as n goes to infinity.

$$\begin{aligned} S &= \lim_{n\to\infty}\sum_{i=1}^{n} 2\pi f(x_i^*)\sqrt{1+\left[f'(x_i^*)\right]^2}\ \ \Delta x \\ &= \int_a^b 2\pi f(x)\sqrt{1+\left[f'(x)\right]^2}\,dx \end{aligned}$$

If we wanted to we could also derive a similar formula for rotating $x = h(y)$ on $[c,d]$ about the y-axis. This would give the following formula.

$$S = \int_c^d 2\pi\, h(y)\sqrt{1+\left[h'(y)\right]^2}\,dy$$

These are not the "standard" formulas however. Notice that the roots in both of these formulas are nothing more than the two *ds*'s we used in the previous section. Also, we will replace $f(x)$ with y and $h(y)$ with x. Doing this gives the following two formulas for the surface area.

Surface Area Formulas

$$S = \int 2\pi y\,ds \qquad \text{rotation about } x-\text{axis}$$

$$S = \int 2\pi x\,ds \qquad \text{rotation about } y-\text{axis}$$

where,

$$ds = \sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx \qquad \text{if } y = f(x),\ a \le x \le b$$

$$ds = \sqrt{1+\left(\frac{dx}{dy}\right)^2}\,dy \qquad \text{if } x = h(y),\ c \le y \le d$$

There are a couple of things to note about these formulas. First, notice that the variable in the integral itself is always the opposite variable from the one we're rotating about. Second, we are allowed to use either *ds* in either formula. This means that there are, in some way, four formulas here. We will choose the *ds* based upon which is the most convenient for a given function and problem.

Now let's work a couple of examples.

Example 1 Determine the surface area of the solid obtained by rotating $y = \sqrt{9 - x^2}$, $-2 \le x \le 2$ about the *x*-axis.

Solution
The formula that we'll be using here is,

$$S = \int 2\pi y \, ds$$

since we are rotating about the *x*-axis and we'll use the first *ds* in this case because our function is in the correct form for that *ds* and we won't gain anything by solving it for *x*.

Let's first get the derivative and the root taken care of.

$$\frac{dy}{dx} = \frac{1}{2}\left(9 - x^2\right)^{-\frac{1}{2}}\left(-2x\right) = -\frac{x}{\left(9 - x^2\right)^{\frac{1}{2}}}$$

$$\sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \sqrt{1 + \frac{x^2}{9 - x^2}} = \sqrt{\frac{9}{9 - x^2}} = \frac{3}{\sqrt{9 - x^2}}$$

Here's the integral for the surface area,

$$S = \int_{-2}^{2} 2\pi y \frac{3}{\sqrt{9 - x^2}} \, dx$$

There is a problem however. The *dx* means that we shouldn't have any *y*'s in the integral. So, before evaluating the integral we'll need to substitute in for *y* as well.

The surface area is then,

$$\begin{aligned} S &= \int_{-2}^{2} 2\pi \sqrt{9 - x^2} \frac{3}{\sqrt{9 - x^2}} \, dx \\ &= \int_{-2}^{2} 6\pi \, dx \\ &= 24\pi \end{aligned}$$

Previously we made the comment that we could use either *ds* in the surface area formulas. Let's work an example in which using either *ds* won't create integrals that are too difficult to evaluate and so we can check both *ds*'s.

Example 2 Determine the surface area of the solid obtained by rotating $y = \sqrt[3]{x}$, $1 \le y \le 2$ about the *y*-axis. Use both *ds*'s to compute the surface area.

Solution
Note that we've been given the function set up for the first *ds* and limits that work for the second *ds*.

Solution 1
This solution will use the first *ds* listed above. We'll start with the derivative and root.

$$\frac{dy}{dx} = \frac{1}{3} x^{-\frac{2}{3}}$$

$$\sqrt{1+\left(\frac{dy}{dx}\right)^2} = \sqrt{1+\frac{1}{9x^{\frac{4}{3}}}} = \sqrt{\frac{9x^{\frac{4}{3}}+1}{9x^{\frac{4}{3}}}} = \frac{\sqrt{9x^{\frac{4}{3}}+1}}{3x^{\frac{2}{3}}}$$

We'll also need to get new limits. That isn't too bad however. All we need to do is plug in the given *y*'s into our equation and solve to get that the range of *x*'s is $1 \le x \le 8$. The integral for the surface area is then,

$$\begin{aligned} S &= \int_1^8 2\pi x \frac{\sqrt{9x^{\frac{4}{3}}+1}}{3x^{\frac{2}{3}}}\, dx \\ &= \frac{2\pi}{3}\int_1^8 x^{\frac{1}{3}}\sqrt{9x^{\frac{4}{3}}+1}\, dx \end{aligned}$$

Note that this time we didn't need to substitute in for the *x* as we did in the previous example. In this case we picked up a *dx* from the *ds* and so we don't need to do a substitution for the *x*. In fact if we had substituted for *x* we would have put *y*'s into the integral which would have caused problems.

Using the substitution

$$u = 9x^{\frac{4}{3}} + 1 \qquad\qquad du = 12x^{\frac{1}{3}}\, dx$$

the integral becomes,

$$\begin{aligned} S &= \frac{\pi}{18}\int_{10}^{145} \sqrt{u}\, du \\ &= \frac{\pi}{27} u^{\frac{3}{2}}\Big|_{10}^{145} \\ &= \frac{\pi}{27}\left(145^{\frac{3}{2}} - 10^{\frac{3}{2}}\right) = 199.48 \end{aligned}$$

Solution 2

This time we'll use the second *ds*. So, we'll first need to solve the equation for *x*. We'll also go ahead and get the derivative and root while we're at it.

$$x = y^3 \qquad\qquad \frac{dx}{dy} = 3y^2$$

$$\sqrt{1+\left(\frac{dx}{dy}\right)^2} = \sqrt{1+9y^4}$$

The surface area is then,

$$S = \int_1^2 2\pi x\sqrt{1+9y^4}\,dy$$

We used the original *y* limits this time because we picked up a *dy* from the *ds*. Also note that the presence of the *dy* means that this time, unlike the first solution, we'll need to substitute in for the *x*. Doing that gives,

$$\begin{aligned} S &= \int_1^2 2\pi y^3\sqrt{1+9y^4}\,dy & u &= 1+9y^4 \\ &= \frac{\pi}{18}\int_{10}^{145}\sqrt{u}\,du \\ &= \frac{\pi}{27}\left(145^{\frac{3}{2}} - 10^{\frac{3}{2}}\right) = 199.48 \end{aligned}$$

Note that after the substitution the integral was identical to the first solution and so the work was skipped.

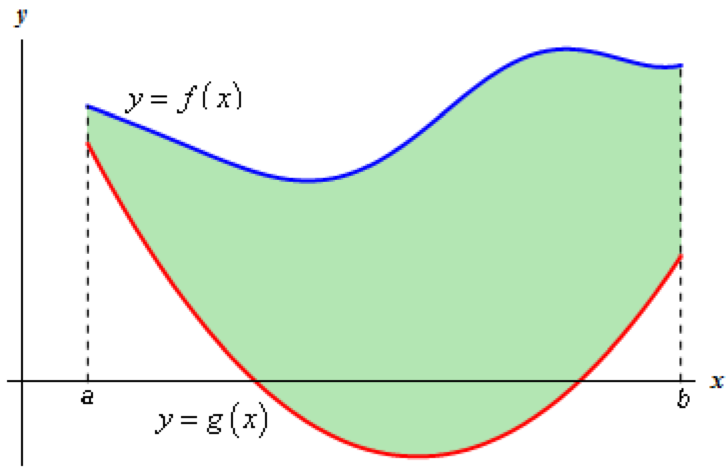
As this example has shown we can use either *ds* to get the surface area. It is important to point out as well that with one *ds* we had to do a substitution for the *x* and with the other we didn't. This will always work out that way.

Note as well that in the case of the last example it was just as easy to use either *ds*. That often won't be the case. In many examples only one of the *ds* will be convenient to work with so we'll always need to determine which *ds* is liable to be the easiest to work with before starting the problem.

Center of Mass

In this section we are going to find the **center of mass** or **centroid** of a thin plate with uniform density ρ. The center of mass or centroid of a region is the point in which the region will be perfectly balanced horizontally if suspended from that point.

So, let's suppose that the plate is the region bounded by the two curves $f(x)$ and $g(x)$ on the interval [a,b]. So, we want to find the center of mass of the region below.



We'll first need the mass of this plate. The mass is,

$$\begin{aligned} M &= \rho\left(\text{Area of plate}\right) \\ &= \rho \int_a^b f(x) - g(x)\,dx \end{aligned}$$

Next we'll need the **moments** of the region. There are two moments, denoted by M_x and M_y. The moments measure the tendency of the region to rotate about the x and y-axis respectively. The moments are given by,

Equations of Moments

$$M_x = \rho \int_a^b \frac{1}{2}\left(\left[f(x)\right]^2 - \left[g(x)\right]^2\right)dx$$

$$M_y = \rho \int_a^b x\left(f(x) - g(x)\right)dx$$

The coordinates of the center of mass, $\left(\overline{x}, \overline{y}\right)$, are then,

Center of Mass Coordinates

$$\bar{x} = \frac{M_y}{M} = \frac{\int_a^b x\left(f(x) - g(x)\right)dx}{\int_a^b f(x) - g(x)\,dx} = \frac{1}{A}\int_a^b x\left(f(x) - g(x)\right)dx$$

$$\bar{y} = \frac{M_x}{M} = \frac{\int_a^b \frac{1}{2}\left(\left[f(x)\right]^2 - \left[g(x)\right]^2\right)dx}{\int_a^b f(x) - g(x)\,dx} = \frac{1}{A}\int_a^b \frac{1}{2}\left(\left[f(x)\right]^2 - \left[g(x)\right]^2\right)dx$$

where,

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

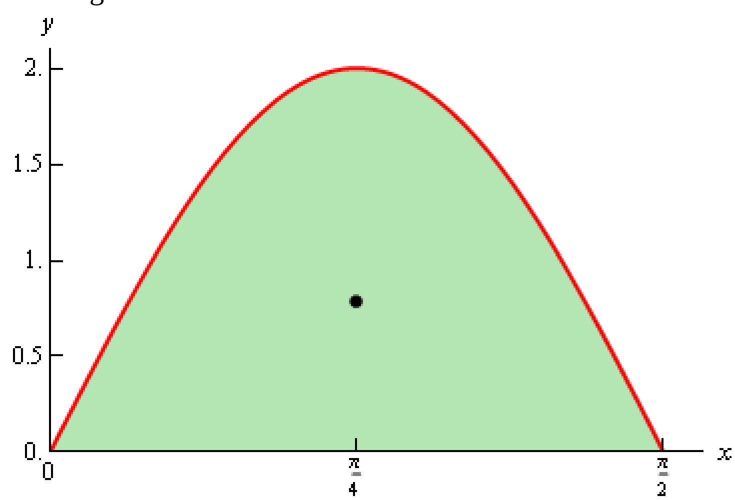
Note that the density, ρ, of the plate cancels out and so isn't really needed.

Let's work a couple of examples.

Example 1 Determine the center of mass for the region bounded by $y = 2\sin(2x)$, $y = 0$ on the interval $\left[0, \frac{\pi}{2}\right]$.

Solution

Here is a sketch of the region with the center of mass denoted with a dot.



Let's first get the area of the region.

$$\begin{aligned} A &= \int_0^{\frac{\pi}{2}} 2\sin(2x)\,dx \\ &= -\cos(2x)\Big|_0^{\frac{\pi}{2}} \\ &= 2 \end{aligned}$$

Now, the moments (without density since it will just drop out) are,

$$M_x = \int_0^{\frac{\pi}{2}} 2\sin^2(2x)\,dx \qquad\qquad M_y = \int_0^{\frac{\pi}{2}} 2x\sin(2x)\,dx \qquad \text{integrating by parts...}$$

$$= \int_0^{\frac{\pi}{2}} 1-\cos(4x)\,dx \qquad\qquad = -x\cos(2x)\Big|_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos(2x)\,dx$$

$$= \left(x - \frac{1}{4}\sin(4x)\right)\Bigg|_0^{\frac{\pi}{2}} \qquad\qquad = -x\cos(2x)\Big|_0^{\frac{\pi}{2}} + \frac{1}{2}\sin(2x)\Bigg|_0^{\frac{\pi}{2}}$$

$$= \frac{\pi}{2} \qquad\qquad = \frac{\pi}{2}$$

The coordinates of the center of mass are then,

$$\overline{x} = \frac{\pi/2}{2} = \frac{\pi}{4}$$

$$\overline{y} = \frac{\pi/2}{2} = \frac{\pi}{4}$$

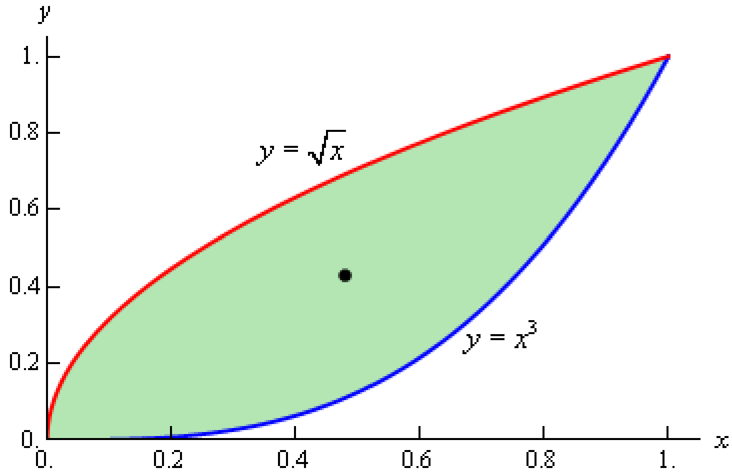
Again, note that we didn't put in the density since it will cancel out.

So, the center of mass for this region is $\left(\frac{\pi}{4}, \frac{\pi}{4}\right)$.

Example 2 Determine the center of mass for the region bounded by $y = x^3$ and $y = \sqrt{x}$.

Solution

The two curves intersect at $x = 0$ and $x = 1$ and here is a sketch of the region with the center of mass marked with a box.



We'll first get the area of the region.

$$A = \int_0^1 \sqrt{x} - x^3 \, dx$$
$$= \left(\frac{2}{3} x^{\frac{3}{2}} - \frac{1}{4} x^4 \right) \Bigg|_0^1$$
$$= \frac{5}{12}$$

Now the moments, again without density, are

$$M_x = \int_0^1 \frac{1}{2}\left(x - x^6 \right) dx$$
$$= \frac{1}{2}\left(\frac{1}{2} x^2 - \frac{1}{7} x^7 \right) \Bigg|_0^1$$
$$= \frac{5}{28}$$

$$M_y = \int_0^1 x\left(\sqrt{x} - x^3 \right) dx$$
$$= \int_0^1 x^{\frac{3}{2}} - x^4 \, dx$$
$$= \left(\frac{2}{5} x^{\frac{5}{2}} - \frac{1}{5} x^5 \right) \Bigg|_0^1$$
$$= \frac{1}{5}$$

The coordinates of the center of mass is then,

$$\overline{x} = \frac{1/5}{5/12} = \frac{12}{25}$$
$$\overline{y} = \frac{5/28}{5/12} = \frac{3}{7}$$

The coordinates of the center of mass are then, $\left(\frac{12}{25}, \frac{3}{7} \right)$.

Hydrostatic Pressure and Force

In this section we are going to submerge a vertical plate in water and we want to know the force that is exerted on the plate due to the pressure of the water. This force is often called the hydrostatic force.

There are two basic formulas that we'll be using here. First, if we are *d* meters below the surface then the hydrostatic pressure is given by,

$$P = \rho g d$$

where, ρ is the density of the fluid and g is the gravitational acceleration. We are going to assume that the fluid in question is water and since we are going to be using the metric system these quantities become,

$$\rho = 1000 \text{ kg/m}^3 \qquad\qquad g = 9.81 \text{ m/s}^2$$

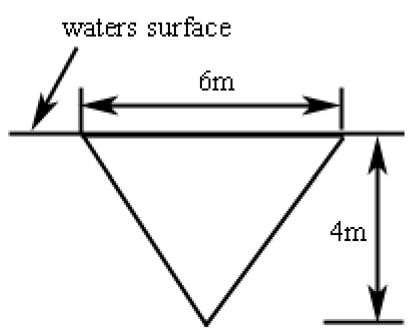
The second formula that we need is the following. Assume that a constant pressure P is acting on a surface with area A. Then the hydrostatic force that acts on the area is,

$$F = PA$$

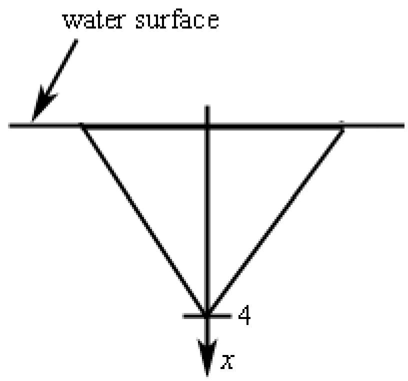
Note that we won't be able to find the hydrostatic force on a vertical plate using this formula since the pressure will vary with depth and hence will not be constant as required by this formula. We will however need this for our work.

The best way to see how these problems work is to do an example or two.

Example 1 Determine the hydrostatic force on the following triangular plate that is submerged in water as shown.



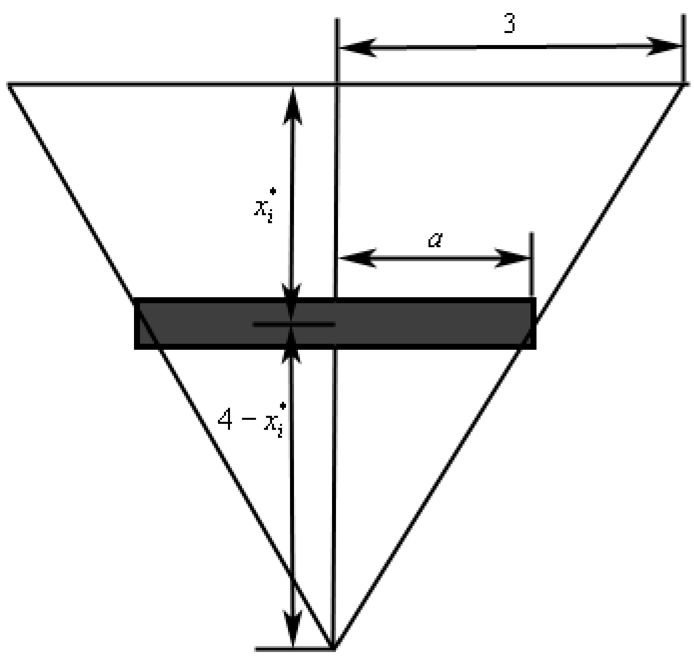
Solution
The first thing to do here is set up an axis system. So, let's redo the sketch above with the following axis system added in.



So, we are going to orient the *x*-axis so that positive *x* is downward, $x = 0$ corresponds to the water surface and $x = 4$ corresponds to the depth of the tip of the triangle.

Next we break up the triangle into *n* horizontal strips each of equal width Δx and in each interval $\left[x_{i-1}, x_i\right]$ choose any point x_i^*. In order to make the computations easier we are going to make two assumptions about these strips. First, we will ignore the fact that the ends are actually going to be slanted and assume the strips are rectangular. If Δx is sufficiently small this will not affect our computations much. Second, we will assume that Δx is small enough that the hydrostatic pressure on each strip is essentially constant.

Below is a representative strip.



The height of this strip is Δx and the width is 2*a*. We can use similar triangles to determine *a* as follows,

$$\frac{3}{4} = \frac{a}{4 - x_i^*} \qquad \Rightarrow \qquad a = 3 - \frac{3}{4} x_i^*$$

Now, since we are assuming the pressure on this strip is constant, the pressure is given by,

$$P_i = \rho g d = 1000(9.81)\,x_i^* = 9810 x_i^*$$

and the hydrostatic force on each strip is,

$$F_i = P_i\,A = P_i\left(2a\Delta x\right) = 9810 x_i^*\,(2)\left(3 - \frac{3}{4}x_i^*\right)\Delta x = 19620 x_i^*\left(3 - \frac{3}{4}x_i^*\right)\Delta x$$

The approximate hydrostatic force on the plate is then the sum of the forces on all the strips or,

$$F \approx \sum_{i=1}^{n} 19620 x_i^*\left(3 - \frac{3}{4}x_i^*\right)\Delta x$$

Taking the limit will get the exact hydrostatic force,

$$F = \lim_{n\to\infty}\sum_{i=1}^{n} 19620 x_i^*\left(3 - \frac{3}{4}x_i^*\right)\Delta x$$

Using the definition of the definite integral this is nothing more than,

$$F = \int_0^4 19620\left(3x - \frac{3}{4}x^2\right)dx$$

The hydrostatic force is then,

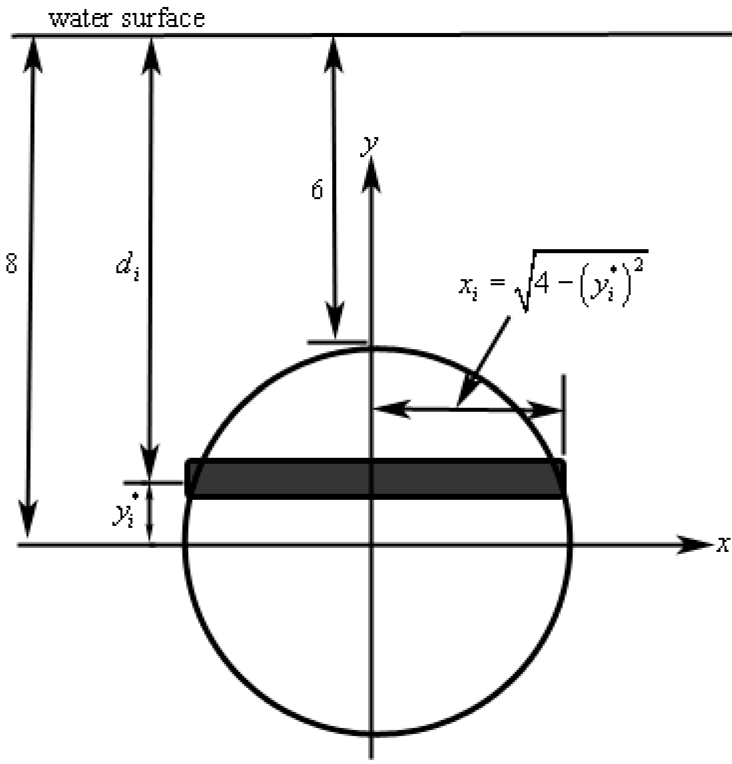
$$\begin{aligned} F &= \int_0^4 19620\left(3x - \frac{3}{4}x^2\right)dx \\ &= 19620\left(\frac{3}{2}x^2 - \frac{1}{4}x^3\right)\Big|_0^4 \\ &= 156960\,N \end{aligned}$$

Let’s take a look at another example.

Example 2 Find the hydrostatic force on a circular plate of radius 2 that is submerged 6 meters in the water.

Solution
First, we’re going to assume that the top of the circular plate is 6 meters under the water. Next, we will set up the axis system so that the origin of the axis system is at the center of the plate. Setting the axis system up in this way will greatly simplify our work.

Finally, we will again split up the plate into *n* horizontal strips each of width Δy and we’ll choose a point y_i^* from each strip. We’ll also assume that the strips are rectangular again to help with the computations. Here is a sketch of the setup.



The depth below the water surface of each strip is,

$$d_i = 8 - y_i^*$$

and that in turn gives us the pressure on the strip,

$$P_i = \rho g d_i = 9810\left(8 - y_i^*\right)$$

The area of each strip is,

$$A_i = 2\sqrt{4-\left(y_i^*\right)^2}\ \Delta y$$

The hydrostatic force on each strip is,

$$F_i = P_i A_i = 9810\left(8 - y_i^*\right)(2)\sqrt{4-\left(y_i^*\right)^2}\ \Delta y$$

The total force on the plate is,

$$\begin{aligned} F &= \lim_{n\to\infty}\sum_{i=1}^{n} 19620\left(8 - y_i^*\right)\sqrt{4-\left(y_i^*\right)^2}\ \Delta y \\ &= 19620\int_{-2}^{2}\left(8-y\right)\sqrt{4-y^2}\,dy \end{aligned}$$

To do this integral we'll need to split it up into two integrals.

$$F = 19620\int_{-2}^{2} 8\sqrt{4-y^2}\,dy - 19620\int_{-2}^{2} y\sqrt{4-y^2}\,dy$$

The first integral requires the trig substitution $y = 2\sin\theta$ and the second integral needs the substitution $v = 4 - y^2$. After using these substitution we get,

$$\begin{aligned} F &= 627840\int_{-\pi/2}^{\pi/2} \cos^2\theta \, d\theta + 9810\int_0^0 \sqrt{v}\, dv \\ &= 313920\int_{-\pi/2}^{\pi/2} 1 + \cos\left(2\theta\right) d\theta + 0 \\ &= 313920\left(\theta + \frac{1}{2}\sin\left(2\theta\right)\right)\Bigg|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\ &= 313920\pi \end{aligned}$$

Note that after the substitution we know the second integral will be zero because the upper and lower limit is the same.

Probability

In this last application of integrals that we'll be looking at we're going to look at probability. Before actually getting into the applications we need to get a couple of definitions out of the way.

Suppose that we wanted to look at the age of a person, the height of a person, the amount of time spent waiting in line, or maybe the lifetime of a battery. Each of these quantities have values that will range over an interval of integers. Because of this these are called **continuous random variables**. Continuous random variables are often represented by X.

Every continuous random variable, X, has a **probability density function**, $f(x)$. Probability density functions satisfy the following conditions.

1. $f(x) \geq 0$ for all x.
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Probability density functions can be used to determine the probability that a continuous random variable lies between two values, say a and b. This probability is denoted by $P(a \leq X \leq b)$ and is given by,

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

Let's take a look at an example of this.

Example 1 Let $f(x) = \dfrac{x^3}{5000}(10 - x)$ for $0 \leq x \leq 10$ and $f(x) = 0$ for all other values of x. Answer each of the following questions about this function.

(a) Show that $f(x)$ is a probability density function. [Solution]

(b) Find $P(1 \leq X \leq 4)$ [Solution]

(c) Find $P(x \geq 6)$ [Solution]

Solution

(a) Show that $f(x)$ is a probability density function.

First note that in the range $0 \leq x \leq 10$ is clearly positive and outside of this range we've defined it to be zero.

So, to show this is a probability density function we'll need to show that $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{10} \frac{x^3}{5000}(10-x)\,dx$$
$$= \left(\frac{x^4}{2000} - \frac{x^5}{25000}\right)\Bigg|_0^{10}$$
$$= 1$$

Note the change in limits on the integral. The function is only non-zero in these ranges and so the integral can be reduced down to only the interval where the function is not zero.

[Return to Problems]

(b) Find $P(1 \le X \le 4)$

In this case we need to evaluate the following integral.

$$P(1 \le X \le 4) = \int_1^4 \frac{x^3}{5000}(10-x)\,dx$$
$$= \left(\frac{x^4}{2000} - \frac{x^5}{25000}\right)\Bigg|_1^4$$
$$= 0.08658$$

So the probability of X being between 1 and 4 is 8.658%.

[Return to Problems]

(c) Find $P(x \ge 6)$

Note that in this case $P(x \ge 6)$ is equivalent to $P(6 \le X \le 10)$ since 10 is the largest value that X can be. So the probability that X is greater than or equal to 6 is,

$$P(X \ge 6) = \int_6^{10} \frac{x^3}{5000}(10-x)\,dx$$
$$= \left(\frac{x^4}{2000} - \frac{x^5}{25000}\right)\Bigg|_6^{10}$$
$$= 0.66304$$

This probability is then 66.304%.

[Return to Problems]

Probability density functions can also be used to determine the mean of a continuous random variable. The mean is given by,

$$\mu = \int_{-\infty}^{\infty} x f(x)\,dx$$

Let's work one more example.

Example 2 It has been determined that the probability density function for the wait in line at a counter is given by,

$$f(t) = \begin{cases} 0 & \text{if } t < 0 \\ 0.1\mathbf{e}^{-\frac{t}{10}} & \text{if } t \geq 0 \end{cases}$$

where t is the number of minutes spent waiting in line. Answer each of the following questions about this probability density function.

(a) Verify that this is in fact a probability density function. [Solution]

(b) Determine the probability that a person will wait in line for at least 6 minutes.\ [Solution]

(c) Determine the mean wait in line. [Solution]

Solution

(a) Verify that this is in fact a probability density function.

This function is clearly positive or zero and so there's not much to do here other than compute the integral.

$$\begin{aligned}\int_{-\infty}^{\infty} f(t)\,dt &= \int_0^{\infty} 0.1\mathbf{e}^{-\frac{t}{10}}\,dt \\ &= \lim_{u\to\infty}\int_0^{u} 0.1\mathbf{e}^{-\frac{t}{10}}\,dt \\ &= \lim_{u\to\infty}\left.\left(-\mathbf{e}^{-\frac{t}{10}}\right)\right|_0^u \\ &= \lim_{u\to\infty}\left(1-\mathbf{e}^{-\frac{u}{10}}\right) = 1\end{aligned}$$

So it is a probability density function.

[Return to Problems]

(b) Determine the probability that a person will wait in line for at least 6 minutes.

The probability that we're looking for here is $P(x \geq 6)$.

$$\begin{aligned}P(X \geq 6) &= \int_6^{\infty} 0.1\mathbf{e}^{-\frac{t}{10}}\,dt \\ &= \lim_{u\to\infty}\int_6^{u} 0.1\mathbf{e}^{-\frac{t}{10}}\,dt \\ &= \lim_{u\to\infty}\left.\left(-\mathbf{e}^{-\frac{t}{10}}\right)\right|_6^u \\ &= \lim_{u\to\infty}\left(\mathbf{e}^{-\frac{6}{10}}-\mathbf{e}^{-\frac{u}{10}}\right) = \mathbf{e}^{-\frac{3}{5}} = 0.548812\end{aligned}$$

So the probability that a person will wait in line for more than 6 minutes is 54.8811%.

[Return to Problems]

(c) Determine the mean wait in line.

Here's the mean wait time.

$$\begin{aligned}\mu &= \int_{-\infty}^{\infty} t\, f(t)\, dt \\ &= \int_{0}^{\infty} 0.1t\,\mathbf{e}^{-\frac{t}{10}}\, dt \\ &= \lim_{u\to\infty}\int_{0}^{u} 0.1t\,\mathbf{e}^{-\frac{t}{10}}\, dt \qquad \text{integrating by parts....} \\ &= \lim_{u\to\infty}\left(-(t+10)\,\mathbf{e}^{-\frac{t}{10}}\right)\Bigg|_{0}^{u} \\ &= \lim_{u\to\infty}\left(10-(u+10)\,\mathbf{e}^{-\frac{u}{10}}\right) = 10\end{aligned}$$

So, it looks like the average wait time is 10 minutes.

[Return to Problems]

Parametric Equations and Polar Coordinates

Introduction

In this section we will be looking at parametric equations and polar coordinates. While the two subjects don't appear to have that much in common on the surface we will see that several of the topics in polar coordinates can be done in terms of parametric equations and so in that sense they make a good match in this chapter

We will also be looking at how to do many of the standard calculus topics such as tangents and area in terms of parametric equations and polar coordinates.

Here is a list of topics that we'll be covering in this chapter.

Parametric Equations and Curves – An introduction to parametric equations and parametric curves (*i.e.* graphs of parametric equations)

Tangents with Parametric Equations – Finding tangent lines to parametric curves.

Area with Parametric Equations – Finding the area under a parametric curve.

Arc Length with Parametric Equations – Determining the length of a parametric curve.

Surface Area with Parametric Equations – Here we will determine the surface area of a solid obtained by rotating a parametric curve about an axis.

Polar Coordinates – We'll introduce polar coordinates in this section. We'll look at converting between polar coordinates and Cartesian coordinates as well as some basic graphs in polar coordinates.

Tangents with Polar Coordinates – Finding tangent lines of polar curves.

Area with Polar Coordinates – Finding the area enclosed by a polar curve.

Arc Length with Polar Coordinates – Determining the length of a polar curve.

Surface Area with Polar Coordinates – Here we will determine the surface area of a solid obtained by rotating a polar curve about an axis.

Arc Length and Surface Area Revisited – In this section we will summarize all the arc length and surface area formulas from the last two chapters.

Parametric Equations and Curves

To this point (in both Calculus I and Calculus II) we've looked almost exclusively at functions in the form $y = f(x)$ or $x = h(y)$ and almost all of the formulas that we've developed require that functions be in one of these two forms. The problem is that not all curves or equations that we'd like to look at fall easily into this form.

Take, for example, a circle. It is easy enough to write down the equation of a circle centered at the origin with radius *r*.

$$x^2 + y^2 = r^2$$

However, we will never be able to write the equation of a circle down as a single equation in either of the forms above. Sure we can solve for *x* or *y* as the following two formulas show

$$y = \pm\sqrt{r^2 - x^2} \qquad\qquad x = \pm\sqrt{r^2 - y^2}$$

but there are in fact two functions in each of these. Each formula gives a portion of the circle.

$$\begin{array}{llll} y = \sqrt{r^2 - x^2} & (\text{top}) & x = \sqrt{r^2 - y^2} & (\text{right side}) \\ y = -\sqrt{r^2 - x^2} & (\text{bottom}) & x = -\sqrt{r^2 - y^2} & (\text{left side}) \end{array}$$

Unfortunately we usually are working on the whole circle, or simply can't say that we're going to be working only on one portion of it. Even if we can narrow things down to only one of these portions the function is still often fairly unpleasant to work with.

There are also a great many curves out there that we can't even write down as a single equation in terms of only *x* and *y*. So, to deal with some of these problems we introduce **parametric equations**. Instead of defining *y* in terms of *x* ($y = f(x)$) or *x* in terms of *y* ($x = h(y)$) we define both *x* and *y* in terms of a third variable called a parameter as follows,

$$x = f(t) \qquad\qquad y = g(t)$$

This third variable is usually denoted by *t* (as we did here) but doesn't have to be of course. Sometimes we will restrict the values of *t* that we'll use and at other times we won't. This will often be dependent on the problem and just what we are attempting to do.

Each value of *t* defines a point $(x, y) = (f(t), g(t))$ that we can plot. The collection of points that we get by letting *t* be all possible values is the graph of the parametric equations and is called the **parametric curve**.

To help visualize just what a parametric curve is pretend that we have a big tank of water that is in constant motion and we drop a ping pong ball into the tank. The point $(x, y) = (f(t), g(t))$ will then represent the location of the ping pong ball in the tank at time *t* and the parametric curve will be a trace of all the locations of the ping pong ball. Note that this is not always a correct analogy but it is useful initially to help visualize just what a parametric curve is.

Sketching a parametric curve is not always an easy thing to do. Let's take a look at an example to see one way of sketching a parametric curve. This example will also illustrate why this method is usually not the best.

Example 1 Sketch the parametric curve for the following set of parametric equations.

$$x = t^2 + t \qquad\qquad y = 2t - 1$$

Solution

At this point our only option for sketching a parametric curve is to pick values of *t*, plug them into the parametric equations and then plot the points. So, let's plug in some *t*'s.

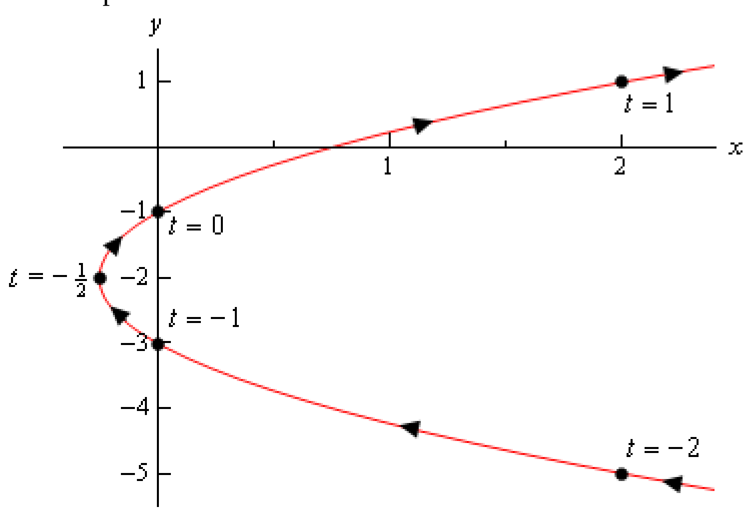
t	x	y
-2	2	-5
-1	0	-3
$-\frac{1}{2}$	$-\frac{1}{4}$	-2
0	0	-1
1	2	1

The first question that should be asked at this point is, how did we know to use the values of *t* that we did, especially the third choice? Unfortunately, there is no real answer to this question at this point. We simply pick *t*'s until we are fairly confident that we've got a good idea of what the curve looks like. It is this problem with picking "good" values of *t* that make this method of sketching parametric curves one of the poorer choices. Sometimes we have no choice, but if we do have a choice we should avoid it.

We'll discuss an alternate graphing method in later examples that will help to explain how these values of *t* were chosen.

We have one more idea to discuss before we actually sketch the curve. Parametric curves have a **direction of motion**. The direction of motion is given by increasing *t*. So, when plotting parametric curves, we also include arrows that show the direction of motion. We will often give the value of *t* that gave specific points on the graph as well to make it clear the value of *t* that gave that particular point.

Here is the sketch of this parametric curve.



So, it looks like we have a parabola that opens to the right.

Before we end this example there is a somewhat important and subtle point that we need to discuss first. Notice that we made sure to include a portion of the sketch to the right of the points corresponding to $t = -2$ and $t = 1$ to indicate that there are portions of the sketch there. Had we simply stopped the sketch at those points we are indicating that there was no portion of the curve to the right of those points and there clearly will be. We just didn't compute any of those points.

This may seem like an unimportant point, but as we'll see in the next example it's more important than we might think.

Before addressing a much easier way to sketch this graph let's first address the issue of limits on the parameter. In the previous example we didn't have any limits on the parameter. Without limits on the parameter the graph will continue in both directions as shown in the sketch above.

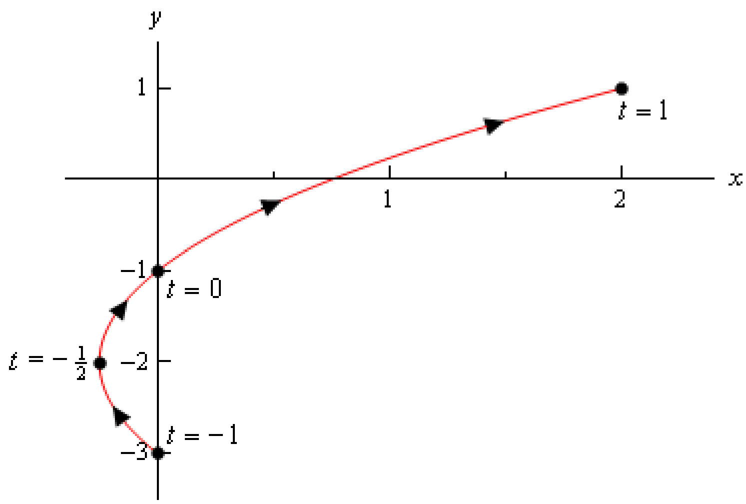
We will often have limits on the parameter however and this will affect the sketch of the parametric equations. To see this effect let's look a slight variation of the previous example.

Example 2 Sketch the parametric curve for the following set of parametric equations.

$$x = t^2 + t \qquad y = 2t - 1 \qquad -1 \le t \le 1$$

Solution

Note that the only difference here is the presence of the limits on *t*. All these limits do is tell us that we can't take any value of *t* outside of this range. Therefore, the parametric curve will only be a portion of the curve above. Here is the parametric curve for this example.



Notice that with this sketch we started and stopped the sketch right on the points originating from the end points of the range of *t*'s. Contrast this with the sketch in the previous example where we had a portion of the sketch to the right of the "start" and "end" points that we computed.

In this case the curve starts at $t = -1$ and ends at $t = 1$, whereas in the previous example the curve didn't really start at the right most points that we computed. We need to be clear in our sketches if the curve starts/ends right at a point, or if that point was simply the first/last one that we computed.

It is now time to take a look at an easier method of sketching this parametric curve. This method uses the fact that in many, but not all, cases we can actually eliminate the parameter from the parametric equations and get a function involving only *x* and *y*. We will sometimes call this the *algebraic equation* to differentiate it from the original parametric equations. There will be two small problems with this method, but it will be easy to address those problems. It is important to note however that we won't always be able to do this.

Just how we eliminate the parameter will depend upon the parametric equations that we've got. Let's see how to eliminate the parameter for the set of parametric equations that we've been working with to this point.

Example 3 Eliminate the parameter from the following set of parametric equations.

$$x = t^2 + t \qquad\qquad y = 2t - 1$$

Solution

One of the easiest ways to eliminate the parameter is to simply solve one of the equations for the parameter (*t*, in this case) and substitute that into the other equation. Note that while this may be the easiest to eliminate the parameter, it's usually not the best way as we'll see soon enough.

In this case we can easily solve *y* for *t*.

$$t = \frac{1}{2}(y+1)$$

Plugging this into the equation for *x* gives the following algebraic equation,

$$x = \left(\frac{1}{2}(y+1)\right)^2 + \frac{1}{2}(y+1) = \frac{1}{4}y^2 + y + \frac{3}{4}$$

Sure enough from our Algebra knowledge we can see that this is a parabola that opens to the right and will have a vertex at $\left(-\frac{1}{4}, -2\right)$.

We won't bother with a sketch for this one as we've already sketched this once and the point here was more to eliminate the parameter anyway.

Before we leave this example let's address one quick issue.

In the first example we just, seemingly randomly, picked values of *t* to use in our table, especially the third value. There really was no apparent reason for choosing $t = -\frac{1}{2}$. It is however probably the most important choice of *t* as it is the one that gives the vertex.

The reality is that when writing this material up we actually did this problem first then went back and did the first problem. Plotting points is generally the way most people first learn how to construct graphs and it does illustrate some important concepts, such as direction, so it made sense to do that first in the notes. In practice however, this example is often done first.

So, how did we get those values of *t*? Well let's start off with the vertex as that is probably the most important point on the graph. We have the *x* and *y* coordinates of the vertex and we also have *x* and *y* parametric equations for those coordinates. So, plug in the coordinates for the vertex into the parametric equations and solve for *t*. Doing this gives,

$$\begin{aligned} -\tfrac{1}{4} &= t^2 + t \\ -2 &= 2t - 1 \end{aligned} \qquad \Rightarrow \qquad \begin{aligned} t &= -\tfrac{1}{2} \quad \text{(double root)} \\ t &= -\tfrac{1}{2} \end{aligned}$$

So, as we can see, the value of t that will give both of these coordinates is $t = -\frac{1}{2}$. Note that the x parametric equation gave a double root and this will often not happen. Often we would have gotten two distinct roots from that equation. In fact, it won't be unusual to get multiple values of t from each of the equations.

However, what we can say is that there will be a value(s) of t that occurs in both sets of solutions and that is the t that we want for that point. We'll eventually see an example where this happens in a later section.

Now, from this work we can see that if we use $t = -\frac{1}{2}$ we will get the vertex and so we included that value of t in the table in Example 1. Once we had that value of t we chose two integer values of t on either side to finish out the table.

As we will see in later examples in this section determining values of t that will give specific points is something that we'll need to do on a fairly regular basis. It is fairly simple however as this example has shown. All we need to be able to do is solve a (usually) fairly basic equation which by this point in time shouldn't be too difficult.

Getting a sketch of the parametric curve once we've eliminated the parameter seems fairly simple. All we need to do is graph the equation that we found by eliminating the parameter. As noted already however, there are two small problems with this method. The first is direction of motion. The equation involving only x and y will NOT give the direction of motion of the parametric curve. This is generally an easy problem to fix however. Let's take a quick look at the derivatives of the parametric equations from the last example. They are,

$$\frac{dx}{dt} = 2t + 1$$
$$\frac{dy}{dt} = 2$$

Now, all we need to do is recall our Calculus I knowledge. The derivative of y with respect to t is clearly always positive. Recalling that one of the interpretations of the first derivative is rate of change we now know that as t increases y must also increase. Therefore, we must be moving up the curve from bottom to top as t increases as that is the only direction that will always give an increasing y as t increases.

Note that the x derivative isn't as useful for this analysis as it will be both positive and negative and hence x will be both increasing and decreasing depending on the value of t. That doesn't help with direction much as following the curve in either direction will exhibit both increasing and decreasing x.

In some cases, only one of the equations, such as this example, will give the direction while in other cases either one could be used. It is also possible that, in some cases, both derivatives would be needed to determine direction. It will always be dependent on the individual set of parametric equations.

The second problem with eliminating the parameter is best illustrated in an example as we'll be running into this problem in the remaining examples.

Example 4 Sketch the parametric curve for the following set of parametric equations. Clearly indicate direction of motion.

$$x = 5\cos t \qquad\qquad y = 2\sin t \qquad\qquad 0 \le t \le 2\pi$$

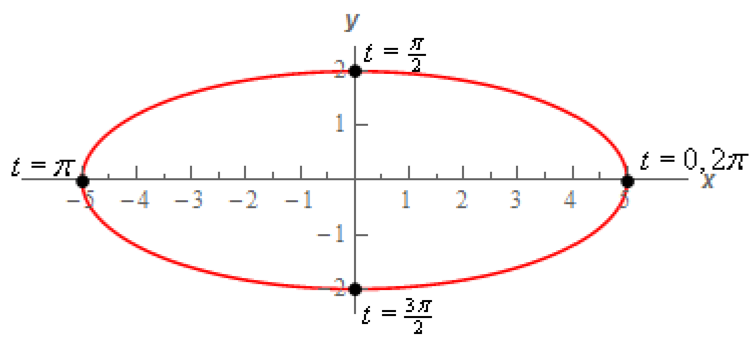
Solution

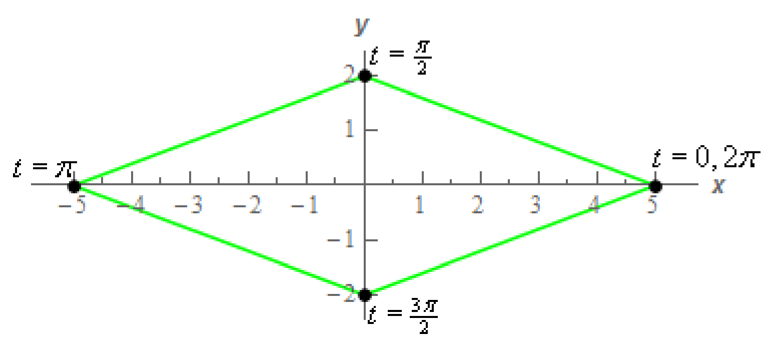
Before we proceed with eliminating the parameter for this problem let's first address again why just picking t's and plotting points is not really a good idea.

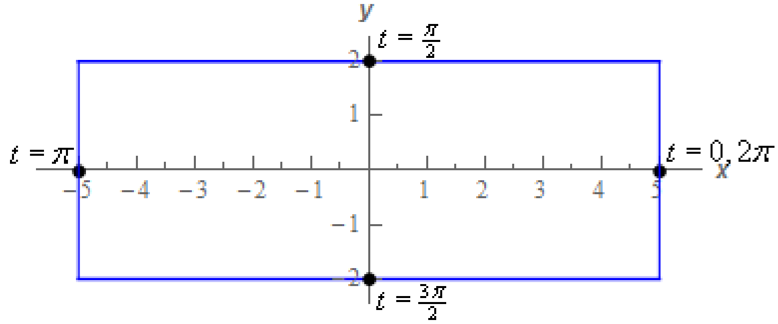
Given the range of t's in the problem statement let's use the following set of t's.

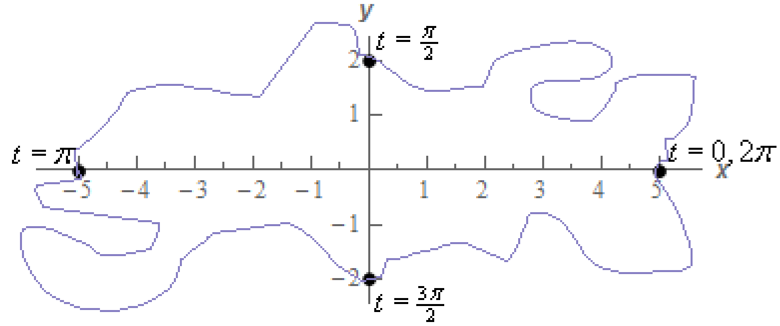
t	x	y
0	5	0
$\frac{\pi}{2}$	0	2
π	-5	0
$\frac{3\pi}{2}$	0	-2
2π	5	0

The question that we need to ask now is do we have enough points to accurately sketch the graph of this set of parametric equations? Below are some sketches of some possible graphs of the parametric equation based only on these five points.









Given the nature of sine/cosine you might be able to eliminate the diamond and the square but there is no denying that they are graphs that go through the given points. The last graph is also a little silly but it does show a graph going through the given points.

Again, given the nature of sine/cosine you can probably guess that the correct graph is the ellipse. However, that is all that would be at this point. A guess. Nothing actually says unequivocally that the parametric curve is an ellipse just from those five points. That is the danger of sketching parametric curves based on a handful of points. Unless we know what the graph will be ahead of time we are really just making a guess.

So, in general, we should avoid plotting points to sketch parametric curves. The best method, provided it can be done, is to eliminate the parameter. As noted just prior to starting this example there is still a potential problem with eliminating the parameter that we'll need to deal with. We will eventually discuss this issue. For now, let's just proceed with eliminating the parameter.

We'll start by eliminating the parameter as we did in the previous section. We'll solve one of the of the equations for t and plug this into the other equation. For example, we could do the following,

$$t = \cos^{-1}\left(\frac{x}{5}\right) \qquad \Rightarrow \qquad y = 2\sin\left(\cos^{-1}\left(\frac{x}{5}\right)\right)$$

Can you see the problem with doing this? This is definitely easy to do but we have a greater chance of correctly graphing the original parametric equations by plotting points than we do graphing this!

There are many ways to eliminate the parameter from the parametric equations and solving for t is usually not the best way to do it. While it is often easy to do we will, in most cases, end up with an equation that is almost impossible to deal with.

So, how can we eliminate the parameter here? In this case all we need to do is recall a very nice trig identity and the equation of an ellipse. Let's notice that we could do the following here.

$$\frac{x^2}{25}+\frac{y^2}{4}=\frac{25\cos^2 t}{25}+\frac{4\sin^2 t}{4}=\cos^2 t+\sin^2 t=1$$

Eliminating the middle steps gives the following algebraic equation,

$$\frac{x^2}{25}+\frac{y^2}{4}=1$$

and so it looks like we've got an ellipse.

Before proceeding with this example it should be noted that what we did was probably not all that obvious to most. However, once it's been done it does clearly work and so it's a nice idea that we can use to eliminate the parameter from some parametric equations involving sines and cosines. It won't always work and sometimes it will take a lot more manipulation of things than we did here.

An alternate method that we could have used here was to solve the two parametric equations for sine and cosine as follows,

$$\cos t=\frac{x}{5} \qquad\qquad \sin t=\frac{y}{2}$$

Then, recall the trig identity we used above and these new equations we get,

$$1=\cos^2 t+\sin^2 t=\left(\frac{x}{5}\right)^2+\left(\frac{y}{2}\right)^2=\frac{x^2}{25}+\frac{y^2}{4}$$

So, the same answer as the other method. Which method you use will probably depend on which you find easier to use. Both are perfectly valid and will get the same result.

Now, let's continue on with the example. We've identified that the parametric equations describe an ellipse, but we can't just sketch an ellipse and be done with it.

First, just because the algebraic equation was an ellipse doesn't actually mean that the parametric curve is the full ellipse. It is always possible that the parametric curve is only a portion of the ellipse. In order to identify just how much of the ellipse the parametric curve will cover let's go back to the parametric equations and see what they tell us about any limits on x and y. Based on our knowledge of sine and cosine we have the following,

$$\begin{array}{lllll} -1\le\cos t\le 1 & \Rightarrow & -5\le 5\cos t\le 5 & \Rightarrow & -5\le x\le 5 \\ -1\le\sin t\le 1 & \Rightarrow & -2\le 2\sin t\le 2 & \Rightarrow & -2\le y\le 2 \end{array}$$

So, by starting with sine/cosine and "building up" the equation for x and y using basic algebraic manipulations we get that the parametric equations enforce the above limits on x and y. In this case, these also happen to be the full limits on x and y we get by graphing the full ellipse.

This is the second potential issue alluded to above. The parametric curve may not always trace out the full graph of the algebraic curve. We should always find limits on *x* and *y* enforced upon us by the parametric curve to determine just how much of the algebraic curve is actually sketched out by the parametric equations.

Therefore, in this case, we now know that we get a full ellipse from the parametric equations. Before we proceed with the rest of the example be careful to not always just assume we will get the full graph of the algebraic equation. There are definitely times when we will not get the full graph and we'll need to do a similar analysis to determine just how much of the graph we actually get. We'll see an example of this later.

Note as well that any limits on *t* given in the problem statement can also affect how much of the graph of the algebraic equation we get. In this case however, based on the table of values we computed at the start of the problem we can see that we do indeed get the full ellipse in the range $0 \le t \le 2\pi$. That won't always be the case however, so pay attention to any restrictions on *t* that might exist!

Next, we need to determine a direction of motion for the parametric curve. Recall that all parametric curves have a direction of motion and the equation of the ellipse simply tells us nothing about the direction of motion.

To get the direction of motion it is tempting to just use the table of values we computed above to get the direction of motion. In this case, we would guess (and yes that is all it is – a guess) that the curve traces out in a counter-clockwise direction. We'd be correct. In this case, we'd be correct! The problem is that tables of values can be misleading when determining a direction of motion as we'll see in the next example.

Therefore, it is best to not use a table of values to determine the direction of motion. To correctly determine the direction of motion we'll use the same method of determining the direction that we discussed after Example 3. In other words, we'll take the derivative of the parametric equations and use our knowledge of Calculus I and trig to determine the direction of motion.

The derivatives of the parametric equations are,

$$\frac{dx}{dt} = -5\sin t \qquad\qquad \frac{dy}{dt} = 2\cos t$$

Now, at $t = 0$ we are at the point $(5,0)$ and let's see what happens if we start increasing *t*. Let's increase *t* from $t = 0$ to $t = \frac{\pi}{2}$. In this range of *t*'s we know that sine is always positive and so from the derivative of the *x* equation we can see that *x* must be decreasing in this range of *t*'s.

This, however, doesn't really help us determine a direction for the parametric curve. Starting at $(5,0)$ no matter if we move in a clockwise or counter-clockwise direction *x* will have to decrease so we haven't really learned anything from the *x* derivative.

The derivative from the *y* parametric equation on the other hand will help us. Again, as we increase *t* from $t = 0$ to $t = \frac{\pi}{2}$ we know that cosine will be positive and so *y* must be increasing in this range. That however, can only happen if we are moving in a counter-clockwise direction.

If we were moving in a clockwise direction from the point $\left(5,0\right)$ we can see that *y* would have to decrease!

Therefore, in the first quadrant we must be moving in a counter-clockwise direction. Let's move on to the second quadrant.

So, we are now at the point $\left(0,2\right)$ and we will increase *t* from $t = \frac{\pi}{2}$ to $t = \pi$. In this range of *t* we know that cosine will be negative and sine will be positive. Therefore, from the derivatives of the parametric equations we can see that *x* is still decreasing and *y* will now be decreasing as well.

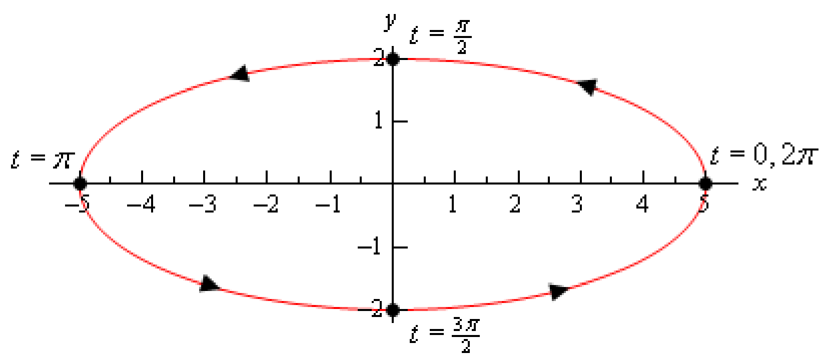
In this quadrant the *y* derivative tells us nothing as *y* simply must decrease to move from $\left(0,2\right)$. However, in order for *x* to decrease, as we know it does in this quadrant, the direction must still be moving a counter-clockwise rotation.

We are now at $\left(-5,0\right)$ and we will increase *t* from $t = \pi$ to $t = \frac{3\pi}{2}$. In this range of *t* we know that cosine is negative (and hence *y* will be decreasing) and sine is also negative (and hence *x* will be increasing). Therefore, we will continue to move in a counter-clockwise motion.

For the 4th quadrant we will start at $\left(0,-2\right)$ and increase *t* from $t = \frac{3\pi}{2}$ to $t = 2\pi$. In this range of *t* we know that cosine is positive (and hence *y* will be increasing) and sine is negative (and hence *x* will be increasing). So, as in the previous three quadrants, we continue to move in a counter-clockwise motion.

At this point we covered the range of *t*'s we were given in the problem statement and during the full range the motion was in a counter-clockwise direction.

We can now fully sketch the parametric curve so, here is the sketch.



Okay, that was a really long example. Most of these types of problems aren't as long. We just had a lot to discuss in this one so we could get a couple of important ideas out of the way. The rest of the examples in this section shouldn't take as long to go through.

Now, let's take a look at another example that will illustrate an important idea about parametric equations.

Example 5 Sketch the parametric curve for the following set of parametric equations. Clearly indicate direction of motion.

$$x = 5\cos(3t) \qquad y = 2\sin(3t) \qquad 0 \le t \le 2\pi$$

Solution

Note that the only difference in between these parametric equations and those in Example 4 is that we replaced the t with $3t$. We can eliminate the parameter here using either of the methods we discussed in the previous example. In this case we'll do the following,

$$\frac{x^2}{25} + \frac{y^2}{4} = \frac{25\cos^2(3t)}{25} + \frac{4\sin^2(3t)}{4} = \cos^2(3t) + \sin^2(3t) = 1$$

So, we get the same ellipse that we did in the previous example. Also note that we can do the same analysis on the parametric equations to determine that we have exactly the same limits on x and y. Namely,

$$-5 \le x \le 5 \qquad -2 \le y \le 2$$

It's starting to look like changing the t into a $3t$ in the trig equations will not change the parametric curve in any way. That is not correct however. The curve does change in a small but important way which we will be discussing shortly.

Before discussing that small change the $3t$ brings to the curve let's discuss the direction of motion for this curve. Despite the fact that we said in the last example that picking values of t and plugging in to the equations to find points to plot is a bad idea let's do it any way.

Given the range of t's from the problem statement the following set looks like a good choice of t's to use.

t	x	y
0	5	0
$\frac{\pi}{2}$	0	-2
π	-5	0
$\frac{3\pi}{2}$	0	2
2π	5	0

So, the only change to this table of values/points from the last example is all the nonzero y values changed sign. From a quick glance at the values in this table it would look like the curve, in this case, is moving in a clockwise direction. But is that correct? Recall we said that these tables of values can be misleading when used to determine direction and that's why we don't use them.

Let's see if our first impression is correct. We can check our first impression by doing the derivative work to get the correct direction. Let's work with just the y parametric equation as the x will have the same issue that it had in the previous example. The derivative of the y parametric equation is,

$$\frac{dy}{dt} = 6\cos(3t)$$

Now, if we start at $t = 0$ as we did in the previous example and start increasing *t*. At $t = 0$ the derivative is clearly positive and so increasing *t* (at least initially) will force *y* to also be increasing. The only way for this to happen is if the curve is in fact tracing out in a counter-clockwise direction initially.

Now, we could continue to look at what happens as we further increase *t*, but when dealing with a parametric curve that is a full ellipse (as this one is) and the argument of the trig functions is of the form *nt* for any constant *n* the direction will not change so once we know the initial direction we know that it will always move in that direction. Note that this is only true for parametric equations in the form that we have here. We'll see in later examples that for different kinds of parametric equations this may no longer be true.

Okay, from this analysis we can see that the curve must be traced out in a counter-clockwise direction. This is directly counter to our guess from the tables of values above and so we can see that, in this case, the table would probably have led us to the wrong direction. So, once again, tables are generally not very reliable for getting pretty much any real information about a parametric curve other than a few points that must be on the curve. Outside of that the tables are rarely useful and will generally not be dealt with in further examples.

So, why did our table give an incorrect impression about the direction? Well recall that we mentioned earlier that the 3*t* will lead to a small but important change to the curve versus just a *t*? Let's take a look at just what that change is as it will also answer what "went wrong" with our table of values.

Let's start by look at $t = 0$. At $t = 0$ we are at the point $\left(5,0\right)$ and let's ask ourselves what values of *t* put us back at this point. We saw in Example 3 how to determine value(s) of *t* that put us at certain points and the same process will work here with a minor modification.

Instead of looking at both the *x* and *y* equations as we did in that example let's just look at the *x* equation. The reason for this is that we'll note that there are two points on the ellipse that will have a *y* coordinate of zero, $\left(5,0\right)$ and $\left(-5,0\right)$. If we set the *y* coordinate equal to zero we'll find all the *t*'s that are at both of these points when we only want the values of *t* that are at $\left(5,0\right)$.

So, because the *x* coordinate of five will only occur at this point we can simply use the *x* parametric equation to determine the values of *t* that will put us at this point. Doing this gives the following equation and solution,

$$\begin{aligned} 5 &= 5\cos\left(3t\right) \\ 3t &= \cos^{-1}\left(1\right) = 0 + 2\pi n \qquad \rightarrow \qquad t = \tfrac{2}{3}\pi n \quad n = 0, \pm 1, \pm 2, \pm 3, \ldots \end{aligned}$$

Don't forget that when solving a trig equation we need to add on the "$+2\pi n$" where *n* represents the number of full revolutions in the counter-clockwise direction (positive *n*) and clockwise direction (negative *n*) that we rotate from the first solution to get all possible solutions to the equation.

Now, let's plug in a few values of *n* starting at $n = 0$. We don't need negative *n* in this case since all of those would result in negative *t* and those fall outside of the range of *t*'s we were given in the problem statement. The first few values of t are then,

$$\begin{array}{lcl} n=0 & : & t=0 \\ n=1 & : & t=\frac{2\pi}{3} \\ n=2 & : & t=\frac{4\pi}{3} \\ n=3 & : & t=\frac{6\pi}{3}=2\pi \end{array}$$

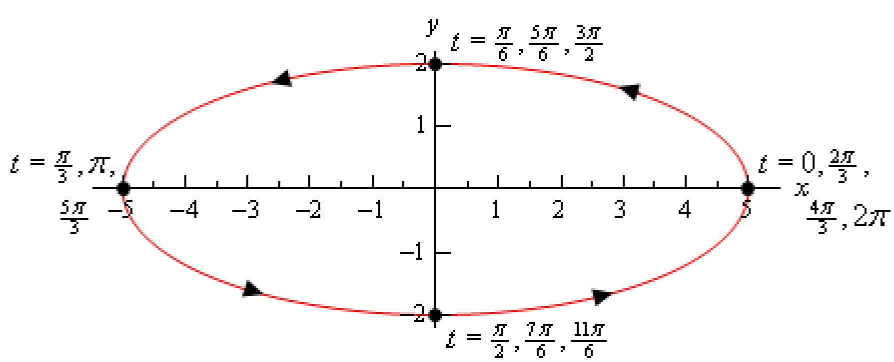
We can stop here as all further values of t will be outside the range of t's given in this problem.

So, what is this telling us? Well back in Example 4 when the argument was just t the ellipse was traced out exactly once in the range $0 \le t \le 2\pi$. However, when we change the argument to $3t$ (and recalling that the curve will always be traced out in a counter-clockwise direction for this problem) we are going through the "starting" point of $\left(5,0\right)$ two more times than we did in the previous example.

In fact, this curve is tracing out three separate times. The first trace is completed in the range $0 \le t \le \frac{2\pi}{3}$. The second trace is completed in the range $\frac{2\pi}{3} \le t \le \frac{4\pi}{3}$ and the third and final trace is completed in the range $\frac{4\pi}{3} \le t \le 2\pi$. In other words, changing the argument from t to $3t$ increase the speed of the trace and the curve will now trace out three times in the range $0 \le t \le 2\pi$!

This is why the table gives the wrong impression. The speed of the tracing has increased leading to an incorrect impression from the points in the table. The table seems to suggest that between each pair of values of t a quarter of the ellipse is traced out in the clockwise direction when in reality it is tracing out three quarters of the ellipse in the counter-clockwise direction.

Here's a final sketch of the curve and note that it really isn't all that different from the previous sketch. The only differences are the values of t and the various points we included. We did include a few more values of t at various points just to illustrate where the curve is at for various values of t but in general these really aren't needed.



So, we saw in the last two examples two sets of parametric equations that in some way gave the same graph. Yet, because they traced out the graph a different number of times we really do need to think of them as different parametric curves at least in some manner. This may seem like a difference that we don't need to worry about, but as we will see in later sections this can be a very important difference. In some of the later sections we are going to need a curve that is traced out exactly once.

Before we move on to other problems let's briefly acknowledge what happens by changing the *t* to an *nt* in these kinds of parametric equations. When we are dealing with parametric equations involving only sines and cosines and they both have the same argument if we change the argument from *t* to *nt* we simply change the speed with which the curve is traced out. If $n > 1$ we will increase the speed and if $n < 1$ we will decrease the speed.

Let's take a look at a couple more examples.

Example 6 Sketch the parametric curve for the following set of parametric equations. Clearly identify the direction of motion. If the curve is traced out more than once give a range of the parameter for which the curve will trace out exactly once.

$$x = \sin^2 t \qquad\qquad y = 2\cos t$$

Solution

We can eliminate the parameter much as we did in the previous two examples. However, we'll need to note that the *x* already contains a $\sin^2 t$ and so we won't need to square the *x*. We will however, need to square the *y* as we need in the previous two examples.

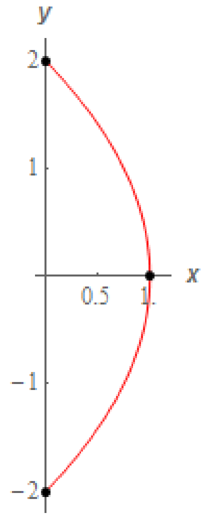
$$x + \frac{y^2}{4} = \sin^2 t + \cos^2 t = 1 \qquad \Rightarrow \qquad x = 1 - \frac{y^2}{4}$$

In this case the algebraic equation is a parabola that opens to the left.

We will need to be very, very careful however in sketching this parametric curve. We will NOT get the whole parabola. A sketch of the algebraic form parabola will exist for all possible values of *y*. However, the parametric equations have defined both *x* and *y* in terms of sine and cosine and we know that the ranges of these are limited and so we won't get all possible values of *x* and *y* here. So, first let's get limits on *x* and *y* as we did in previous examples. Doing this gives,

$$\begin{aligned} -1 \le \sin t \le 1 \qquad &\Rightarrow \qquad 0 \le \sin^2 t \le 1 \qquad \Rightarrow \qquad 0 \le x \le 1 \\ -1 \le \cos t \le 1 \qquad &\Rightarrow \qquad -2 \le 2\cos t \le 2 \qquad \Rightarrow \qquad -2 \le y \le 2 \end{aligned}$$

So, it is clear from this that we will only get a portion of the parabola that is defined by the algebraic equation. Below is a quick sketch of the portion of the parabola that the parametric curve will cover.



To finish the sketch of the parametric curve we also need the direction of motion for the curve. Before we get to that however, let's jump forward and determine the range of *t*'s for one trace. To do this we'll need to know the *t*'s that put us at each end point and we can follow the same procedure we used in the previous example. The only difference is this time let's use the *y* parametric equation instead of the *x* because the *y* coordinates of the two end points of the curve are different whereas the *x* coordinates are the same.

So, for the top point we have,

$$\begin{aligned} 2 &= 2\cos t \\ t &= \cos^{-1}(1) = 0 + 2\pi n = 2\pi n, \qquad n = 0, \pm 1, \pm 2, \pm 3, \ldots \end{aligned}$$

For, plugging in some values of *n* we get that the curve will be at the top point at,

$$t = \ldots, -4\pi, -2\pi, 0, 2\pi, 4\pi, \ldots$$

Similarly, for the bottom point we have,

$$\begin{aligned} -2 &= 2\cos t \\ t &= \cos^{-1}(-1) = \pi + 2\pi n, \qquad n = 0, \pm 1, \pm 2, \pm 3, \ldots \end{aligned}$$

So, we see that we will be at the bottom point at,

$$t = \ldots, -3\pi, -\pi, \pi, 3\pi, \ldots$$

So, if we start at say, $t = 0$, we are at the top point and we increase *t* we have to move along the curve downwards until we reach $t = \pi$ at which point we are now at the bottom point. This means that we will trace out the curve exactly once in the range $0 \le t \le \pi$.

This is not the only range that will trace out the curve however. Note that if we further increase *t* from $t = \pi$ we will now have to travel back up the curve until we reach $t = 2\pi$ and we are now back at the top point. Increasing *t* again until we reach $t = 3\pi$ will take us back down the curve until we reach the bottom point again, *etc*. From this analysis we can get two more ranges of t for one trace,

$$\pi \le t \le 2\pi \qquad\qquad 2\pi \le t \le 3\pi$$

As you can probably see there are an infinite number of ranges of t we could use for one trace of the curve. Any of them would be acceptable answers for this problem.

Note that in the process of determining a range of t's for one trace we also managed to determine the direction of motion for this curve. In the range $0 \le t \le \pi$ we had to travel downwards along the curve to get from the top point at $t = 0$ to the bottom point at $t = \pi$. However, at $t = 2\pi$ we are back at the top point on the curve and to get there we must travel along the path. We can't just jump back up to the top point or take a different path to get there. All travel must be done on the path sketched out. This means that we had to go back up the path. Further increasing t takes us back down the path, then up the path again *etc*.

In other words, this path is sketched out in both directions because we are not putting any restrictions on the t's and so we have to assume we are using all possible values of t. If we had put restrictions on which t's to use we might really have ended up only moving in one direction. That however would be a result only of the range of t's we are using and not the parametric equations themselves.

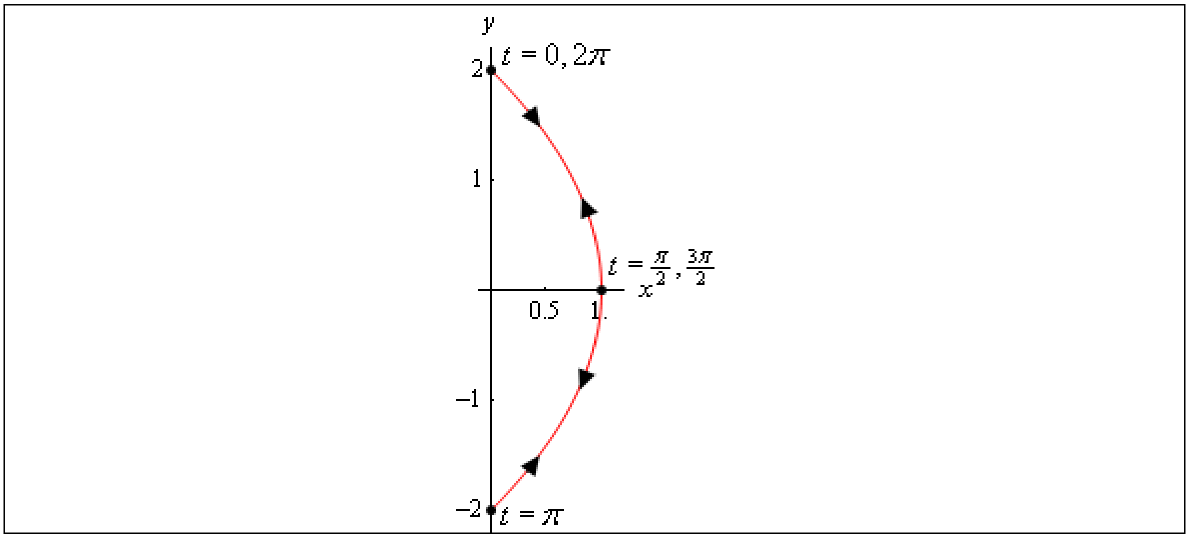
Note that we didn't really need to do the above work to determine that the curve traces out in both directions.in this case. Both the x and y parametric equations involve sine or cosine and we know both of those functions oscillate. This, in turn means that both x and y will oscillate as well. The only way for that to happen on this particular this curve will be for the curve to be traced out in both directions.

Be careful with the above reasoning that the oscillatory nature of sine/cosine forces the curve to be traced out in both directions. It can only be used in this example because the "starting" point and "ending" point of the curves are in different places. The only way to get from one of the "end" points on the curve to the other is to travel back along the curve in the opposite direction.

Contrast this with the ellipse in Example 4. In that case we had sine/cosine in the parametric equations as well. However, the curve only traced out in one direction, not in both directions. In Example 4 we were graphing the full ellipse and so no matter where we start sketching the graph we will eventually get back to the "starting" point without ever retracing any portion of the graph. In Example 4 as we trace out the full ellipse both x and y do in fact oscillate between their two "endpoints" but the curve itself does not trace out in both directions for this to happen.

Basically, we can only use the oscillatory nature of sine/cosine to determine that the curve traces out in both directions if the curve starts and ends at different points. If the starting/ending point is the same then we generally need to go through the full derivative argument to determine the actual direction of motion.

So, to finish this problem out, below is a sketch of the parametric curve. Note that we put direction arrows in both directions to clearly indicate that it would be traced out in both directions. We also put in a few values of t just to help illustrate the direction of motion.



To this point we've seen examples that would trace out the complete graph that we got by eliminating the parameter if we took a large enough range of t's. However, in the previous example we've now seen that this will not always be the case. It is more than possible to have a set of parametric equations which will continuously trace out just a portion of the curve. We can usually determine if this will happen by looking for limits on x and y that are imposed up us by the parametric equation.

We will often use parametric equations to describe the path of an object or particle. Let's take a look at an example of that.

Example 7 The path of a particle is given by the following set of parametric equations.

$$x = 3\cos(2t) \qquad\qquad y = 1 + \cos^2(2t)$$

Completely describe the path of this particle. Do this by sketching the path, determining limits on x and y and giving a range of t's for which the path will be traced out exactly once (provide it traces out more than once of course).

Solution

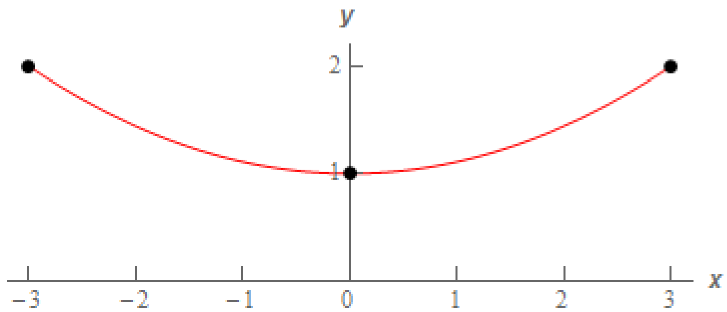
Eliminating the parameter this time will be a little different. We only have cosines this time and we'll use that to our advantage. We can solve the x equation for cosine and plug that into the equation for y. This gives,

$$\cos(2t) = \frac{x}{3} \qquad\qquad y = 1 + \left(\frac{x}{3}\right)^2 = 1 + \frac{x^2}{9}$$

This time the algebraic equation is a parabola that opens upward. We also have the following limits on x and y.

$$\begin{array}{lll} -1 \le \cos(2t) \le 1 & \qquad -3 \le 3\cos(2t) \le 3 & \qquad -3 \le x \le 3 \\ 0 \le \cos^2(2t) \le 1 & \qquad 1 \le 1 + \cos^2(2t) \le 2 & \qquad 1 \le y \le 2 \end{array}$$

So, again we only trace out a portion of the curve. Here is a quick sketch of the portion of the parabola that the parametric curve will cover.



Now, as we discussed in the previous example because both the *x* and *y* parametric equations involve cosine we know that both *x* and *y* must oscillate and because the "start" and "end" points of the curve are not the same the only way *x* and *y* can oscillate is for the curve to trace out in both directions.

To finish the problem then all we need to do is determine a range of *t*'s for one trace. Because the "end" points on the curve have the same *y* value and different *x* values we can use the *x* parametric equation to determine these values. Here is that work.

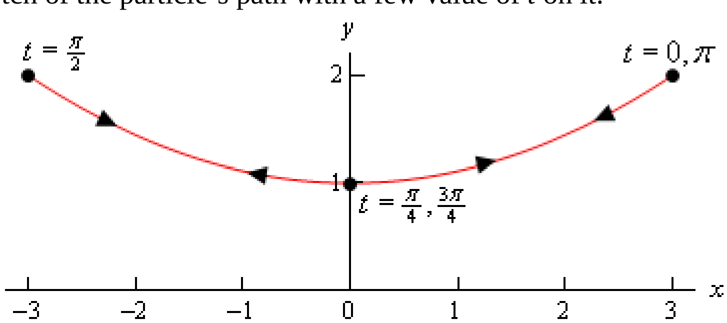
$$x = 3: \qquad 3 = 3\cos(2t)$$
$$1 = \cos(2t)$$
$$2t = 0 + 2\pi n \qquad \rightarrow \qquad t = \pi n \qquad n = 0, \pm1, \pm2, \pm3, \ldots$$

$$x = -3: \qquad -3 = 3\cos(2t)$$
$$-1 = \cos(2t)$$
$$2t = \pi + 2\pi n \qquad \rightarrow \qquad t = \tfrac{1}{2}\pi + \pi n \qquad n = 0, \pm1, \pm2, \pm3, \ldots$$

So, we will be at the right end point at $t = \ldots, -2\pi, -\pi, 0, \pi, 2\pi, \ldots$ and we'll be at the left end point at $t = \ldots, -\tfrac{3}{2}\pi, -\tfrac{1}{2}\pi, \tfrac{1}{2}\pi, \tfrac{3}{2}\pi, \ldots$. So, in this case there are an infinite number of ranges of *t*'s for one trace. Here are a few of them.

$$-\tfrac{1}{2}\pi \le t \le 0 \qquad\qquad 0 \le t \le \tfrac{1}{2}\pi \qquad\qquad \tfrac{1}{2}\pi \le t \le \pi$$

Here is a final sketch of the particle's path with a few value of *t* on it.



We should give a small warning at this point. Because of the ideas involved in them we concentrated on parametric curves that retraced portions of the curve more than once. Do not, however, get too locked into the idea that this will always happen. Many, if not most parametric curves will only trace out once. The first one we looked at is a good example of this. That parametric curve will never repeat any portion of itself.

There is one final topic to be discussed in this section before moving on. So far we've started with parametric equations and eliminated the parameter to determine the parametric curve.

However, there are times in which we want to go the other way. Given a function or equation we might want to write down a set of parametric equations for it. In these cases we say that we **parameterize** the function.

If we take Examples 4 and 5 as examples we can do this for ellipses (and hence circles). Given the ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

a set of parametric equations for it would be,

$$x = a\cos t \qquad\qquad y = b\sin t$$

This set of parametric equations will trace out the ellipse starting at the point $(a,0)$ and will trace in a counter-clockwise direction and will trace out exactly once in the range $0 \le t \le 2\pi$. This is a fairly important set of parametric equations as it used continually in some subjects with dealing with ellipses and/or circles.

Every curve can be parameterized in more than one way. Any of the following will also parameterize the same ellipse.

$$\begin{aligned} x &= a\cos(\omega t) & y &= b\sin(\omega t) \\ x &= a\sin(\omega t) & y &= b\cos(\omega t) \\ x &= a\cos(\omega t) & y &= -b\sin(\omega t) \end{aligned}$$

The presence of the ω will change the speed that the ellipse rotates as we saw in Example 5. Note as well that the last two will trace out ellipses with a clockwise direction of motion (you might want to verify this). Also note that they won't all start at the same place (if we think of $t = 0$ as the starting point that is).

There are many more parameterizations of an ellipse of course, but you get the idea. It is important to remember that each parameterization will trace out the curve once with a potentially different range of *t*'s. Each parameterization may rotate with different directions of motion and may start at different points.

You may find that you need a parameterization of an ellipse that starts at a particular place and has a particular direction of motion and so you now know that with some work you can write down a set of parametric equations that will give you the behavior that you're after.

Now, let's write down a couple of other important parameterizations and all the comments about direction of motion, starting point, and range of *t*'s for one trace (if applicable) are still true.

First, because a circle is nothing more than a special case of an ellipse we can use the parameterization of an ellipse to get the parametric equations for a circle centered at the origin of radius *r* as well. One possible way to parameterize a circle is,

$$x = r\cos t \qquad\qquad y = r\sin t$$

Finally, even though there may not seem to be any reason to, we can also parameterize functions in the form $y = f\left(x\right)$ or $x = h\left(y\right)$. In these cases we parameterize them in the following way,

$$\begin{aligned} x &= t \\ y &= f\left(t\right) \end{aligned} \qquad\qquad \begin{aligned} x &= h\left(t\right) \\ y &= t \end{aligned}$$

At this point it may not seem all that useful to do a parameterization of a function like this, but there are many instances where it will actually be easier, or it may even be required, to work with the parameterization instead of the function itself. Unfortunately, almost all of these instances occur in a Calculus III course.

Tangents with Parametric Equations

In this section we want to find the tangent lines to the parametric equations given by,

$$x = f(t) \qquad\qquad y = g(t)$$

To do this let's first recall how to find the tangent line to $y = F(x)$ at $x = a$. Here the tangent line is given by,

$$y = F(a) + m(x-a), \quad \text{where} \quad m = \left.\frac{dy}{dx}\right|_{x=a} = F'(a)$$

Now, notice that if we could figure out how to get the derivative $\frac{dy}{dx}$ from the parametric equations we could simply reuse this formula since we will be able to use the parametric equations to find the *x* and *y* coordinates of the point.

So, just for a second let's suppose that we were able to eliminate the parameter from the parametric form and write the parametric equations in the form $y = F(x)$. Now, plug the parametric equations in for *x* and *y*. Yes, it seems silly to eliminate the parameter, then immediately put it back in, but it's what we need to do in order to get our hands on the derivative. Doing this gives,

$$g(t) = F\left(f(t)\right)$$

Now, differentiate with respect to *t* and notice that we'll need to use the Chain Rule on the right hand side.

$$g'(t) = F'\left(f(t)\right)\, f'(t)$$

Let's do another change in notation. We need to be careful with our derivatives here. Derivatives of the lower case function are with respect to *t* while derivatives of upper case functions are with respect to *x*. So, to make sure that we keep this straight let's rewrite things as follows.

$$\frac{dy}{dt} = F'(x)\frac{dx}{dt}$$

At this point we should remind ourselves just what we are after. We needed a formula for $\frac{dy}{dx}$ or $F'(x)$ that is in terms of the parametric formulas. Notice however that we can get that from the above equation.

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}}, \qquad \text{provided} \qquad \frac{dx}{dt} \neq 0$$

Notice as well that this will be a function of *t* and not *x*.

As an aside, notice that we could also get the following formula with a similar derivation if we needed to,

Derivative for Parametric Equations

$$\frac{dx}{dy} = \frac{\frac{dx}{dt}}{\frac{dy}{dt}}, \qquad \text{provided} \qquad \frac{dy}{dt} \neq 0$$

Why would we want to do this? Well, recall that in the arc length section of the Applications of Integral section we actually needed this derivative on occasion.

So, let's find a tangent line.

Example 1 Find the tangent line(s) to the parametric curve given by

$$x = t^5 - 4t^3 \qquad\qquad y = t^2$$

at (0,4).

Solution

Note that there is apparently the potential for more than one tangent line here! We will look into this more after we're done with the example.

The first thing that we should do is find the derivative so we can get the slope of the tangent line.

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{2t}{5t^4 - 12t^2} = \frac{2}{5t^3 - 12t}$$

At this point we've got a small problem. The derivative is in terms of *t* and all we've got is an *x*-*y* coordinate pair. The next step then is to determine the value(s) of *t* which will give this point. We find these by plugging the *x* and *y* values into the parametric equations and solving for *t*.

$$0 = t^5 - 4t^3 = t^3\left(t^2 - 4\right) \qquad \Rightarrow \qquad t = 0, \pm 2$$

$$4 = t^2 \qquad \Rightarrow \qquad t = \pm 2$$

Any value of *t* which appears in both lists will give the point. So, since there are two values of *t* that give the point we will in fact get two tangent lines. That's definitely not something that happened back in Calculus I and we're going to need to look into this a little more. However, before we do that let's actually get the tangent lines.

$t = -2$

Since we already know the *x* and *y*-coordinates of the point all that we need to do is find the slope of the tangent line.

$$m = \left.\frac{dy}{dx}\right|_{t=-2} = -\frac{1}{8}$$

The tangent line (at $t = -2$) is then,

$$y = 4 - \frac{1}{8}x$$

$t = 2$
Again, all we need is the slope.

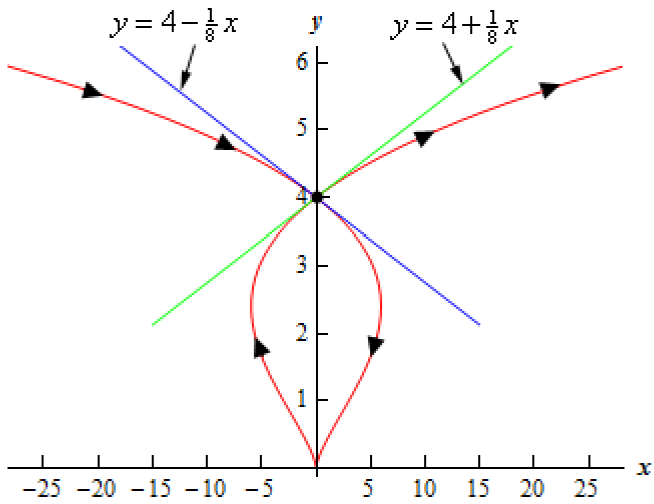
$$m = \left.\frac{dy}{dx}\right|_{t=2} = \frac{1}{8}$$

The tangent line (at $t = 2$) is then,

$$y = 4 + \frac{1}{8}x$$

Now, let's take a look at just how we could possibly get two tangents lines at a point. This was definitely not possible back in Calculus I where we first ran across tangent lines.

A quick graph of the parametric curve will explain what is going on here.



So, the parametric curve crosses itself! That explains how there can be more than one tangent line. There is one tangent line for each instance that the curve goes through the point.

The next topic that we need to discuss in this section is that of horizontal and vertical tangents. We can easily identify where these will occur (or at least the t's that will give them) by looking at the derivative formula.

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

Horizontal tangents will occur where the derivative is zero and that means that we'll get horizontal tangent at values of t for which we have,

Horizontal Tangent for Parametric Equations

$$\frac{dy}{dt} = 0, \quad \text{provided} \quad \frac{dx}{dt} \neq 0$$

Vertical tangents will occur where the derivative is not defined and so we'll get vertical tangents at values of *t* for which we have,

Vertical Tangent for Parametric Equations

$$\frac{dx}{dt} = 0, \quad \text{provided} \quad \frac{dy}{dt} \neq 0$$

Let's take a quick look at an example of this.

Example 2 Determine the *x*-*y* coordinates of the points where the following parametric equations will have horizontal or vertical tangents.

$$x = t^3 - 3t \qquad\qquad y = 3t^2 - 9$$

Solution

We'll first need the derivatives of the parametric equations.

$$\frac{dx}{dt} = 3t^2 - 3 = 3\left(t^2 - 1\right) \qquad\qquad \frac{dy}{dt} = 6t$$

Horizontal Tangents

We'll have horizontal tangents where,

$$6t = 0 \qquad \Rightarrow \qquad t = 0$$

Now, this is the value of *t* which gives the horizontal tangents and we were asked to find the *x*-*y* coordinates of the point. To get these we just need to plug *t* into the parametric equations. Therefore, the only horizontal tangent will occur at the point (0,-9).

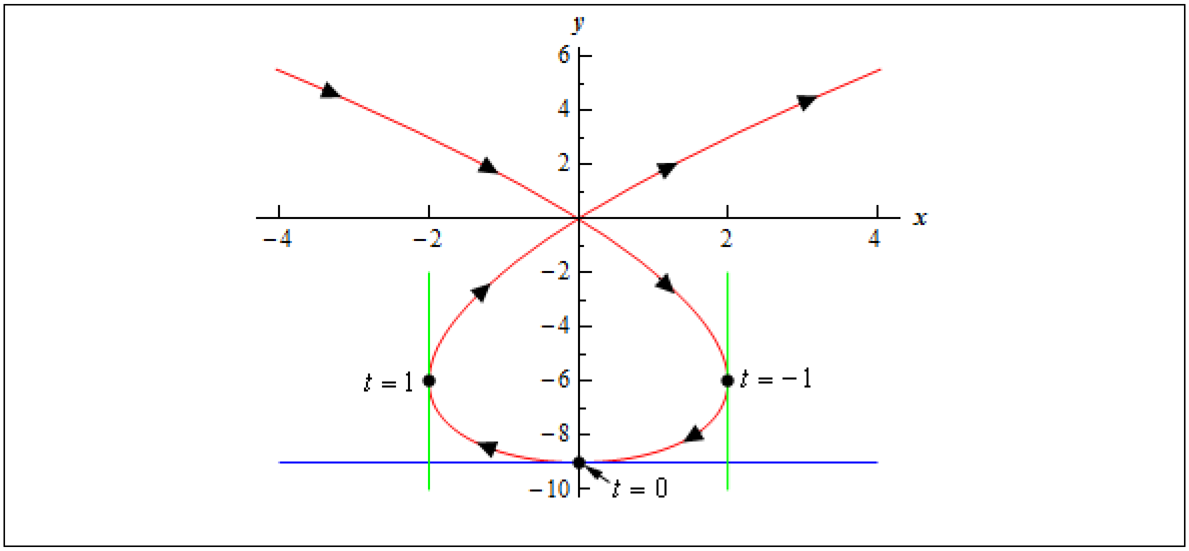
Vertical Tangents

In this case we need to solve,

$$3\left(t^2 - 1\right) = 0 \qquad \Rightarrow \qquad t = \pm 1$$

The two vertical tangents will occur at the points (2,-6) and (-2,-6).

For the sake of completeness and at least partial verification here is the sketch of the parametric curve.



The final topic that we need to discuss in this section really isn't related to tangent lines, but does fit in nicely with the derivation of the derivative that we needed to get the slope of the tangent line.

Before moving into the new topic let's first remind ourselves of the formula for the first derivative and in the process rewrite it slightly.

$$\frac{dy}{dx} = \frac{d}{dx}\left(y\right) = \frac{\frac{d}{dt}\left(y\right)}{\frac{dx}{dt}}$$

Written in this way we can see that the formula actually tells us how to differentiate a function y (as a function of t) with respect to x (when x is also a function of t) when we are using parametric equations.

Now let's move onto the final topic of this section. We would also like to know how to get the second derivative of y with respect to x.

$$\frac{d^2y}{dx^2}$$

Getting a formula for this is fairly simple if we remember the rewritten formula for the first derivative above.

Second Derivative for Parametric Equations

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{\frac{d}{dt}\left(\frac{dy}{dx}\right)}{\frac{dx}{dt}}$$

Note that,

$$\frac{d^2y}{dx^2} \neq \frac{\dfrac{d^2y}{dt^2}}{\dfrac{d^2x}{dt^2}}$$

Let's work a quick example.

Example 3 Find the second derivative for the following set of parametric equations.

$$x = t^5 - 4t^3 \qquad\qquad y = t^2$$

Solution

This is the set of parametric equations that we used in the first example and so we already have the following computations completed.

$$\frac{dy}{dt} = 2t \qquad\qquad \frac{dx}{dt} = 5t^4 - 12t^2 \qquad\qquad \frac{dy}{dx} = \frac{2}{5t^3 - 12t}$$

We will first need the following,

$$\frac{d}{dt}\left(\frac{2}{5t^3 - 12t}\right) = \frac{-2\left(15t^2 - 12\right)}{\left(5t^3 - 12t\right)^2} = \frac{24 - 30t^2}{\left(5t^3 - 12t\right)^2}$$

The second derivative is then,

$$\begin{aligned} \frac{d^2y}{dx^2} &= \frac{\dfrac{d}{dt}\left(\dfrac{dy}{dx}\right)}{\dfrac{dx}{dt}} \\ &= \frac{\dfrac{24 - 30t^2}{\left(5t^3 - 12t\right)^2}}{5t^4 - 12t^2} \\ &= \frac{24 - 30t^2}{\left(5t^4 - 12t^2\right)\left(5t^3 - 12t\right)^2} \\ &= \frac{24 - 30t^2}{t\left(5t^3 - 12t\right)^3} \end{aligned}$$

So, why would we want the second derivative? Well, recall from your Calculus I class that with the second derivative we can determine where a curve is concave up and concave down. We could do the same thing with parametric equations if we wanted to.

Example 4 Determine the values of t for which the parametric curve given by the following set of parametric equations is concave up and concave down.

$$x = 1 - t^2 \qquad\qquad y = t^7 + t^5$$

Solution

To compute the second derivative we'll first need the following.

$$\frac{dy}{dt} = 7t^6 + 5t^4 \qquad\qquad \frac{dx}{dt} = -2t \qquad\qquad \frac{dy}{dx} = \frac{7t^6 + 5t^4}{-2t} = -\frac{1}{2}\left(7t^5 + 5t^3\right)$$

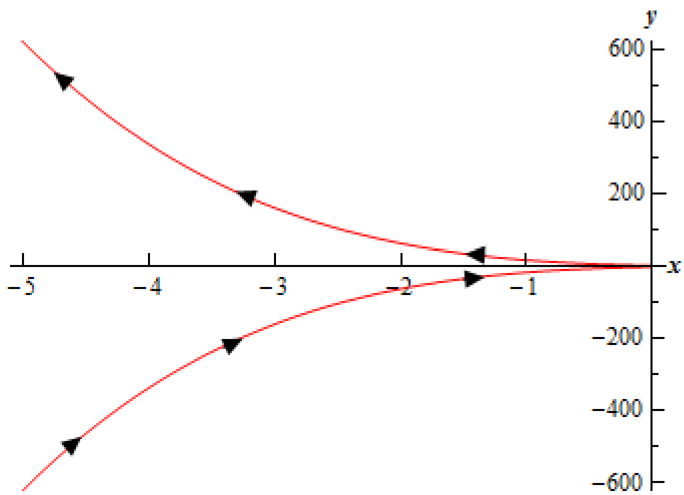
Note that we can also use the first derivative above to get some information about the increasing/decreasing nature of the curve as well. In this case it looks like the parametric curve will be increasing if $t < 0$ and decreasing if $t > 0$.

Now let's move on to the second derivative.

$$\frac{d^2y}{dx^2} = \frac{-\frac{1}{2}\left(35t^4 + 15t^2\right)}{-2t} = \frac{1}{4}\left(35t^3 + 15t\right)$$

It's clear, hopefully, that the second derivative will only be zero at $t = 0$. Using this we can see that the second derivative will be negative if $t < 0$ and positive if $t > 0$. So the parametric curve will be concave down for $t < 0$ and concave up for $t > 0$.

Here is a sketch of the curve for completeness sake.



Area with Parametric Equations

In this section we will find a formula for determining the area under a parametric curve given by the parametric equations,

$$x = f(t) \qquad\qquad y = g(t)$$

We will also need to further add in the assumption that the curve is traced out exactly once as t increases from α to β.

We will do this in much the same way that we found the first derivative in the previous section. We will first recall how to find the area under $y = F(x)$ on $a \le x \le b$.

$$A = \int_a^b F(x)\,dx$$

We will now think of the parametric equation $x = f(t)$ as a substitution in the integral. We will also assume that $a = f(\alpha)$ and $b = f(\beta)$ for the purposes of this formula. There is actually no reason to assume that this will always be the case and so we'll give a corresponding formula later if it's the opposite case ($b = f(\alpha)$ and $a = f(\beta)$).

So, if this is going to be a substitution we'll need,

$$dx = f'(t)\,dt$$

Plugging this into the area formula above and making sure to change the limits to their corresponding t values gives us,

$$A = \int_\alpha^\beta F(f(t))\,f'(t)\,dt$$

Since we don't know what *F(x)* is we'll use the fact that

$$y = F(x) = F(f(t)) = g(t)$$

and we arrive at the formula that we want.

Area Under Parametric Curve, Formula I

$$A = \int_\alpha^\beta g(t)\,f'(t)\,dt$$

Now, if we should happen to have $b = f(\alpha)$ and $a = f(\beta)$ the formula would be,

Area Under Parametric Curve, Formula II

$$A = \int_\beta^\alpha g(t)\,f'(t)\,dt$$

Let's work an example.

Example 1 Determine the area under the parametric curve given by the following parametric equations.

$$x = 6(\theta - \sin\theta) \qquad y = 6(1 - \cos\theta) \qquad 0 \le \theta \le 2\pi$$

Solution

First, notice that we've switched the parameter to θ for this problem. This is to make sure that we don't get too locked into always having *t* as the parameter.

Now, we could graph this to verify that the curve is traced out exactly once for the given range if we wanted to. We are going to be looking at this curve in more detail after this example so we won't sketch its graph here.

There really isn't too much to this example other than plugging the parametric equations into the formula. We'll first need the derivative of the parametric equation for *x* however.

$$\frac{dx}{d\theta} = 6(1 - \cos\theta)$$

The area is then,

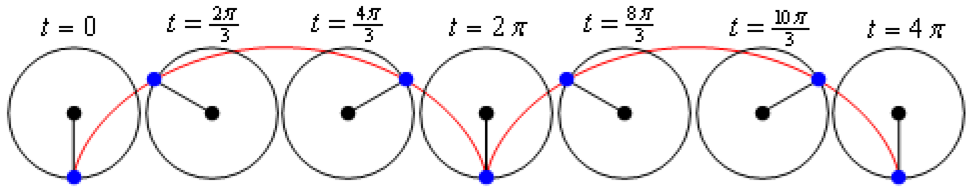
$$\begin{aligned} A &= \int_0^{2\pi} 36(1 - \cos\theta)^2 \, d\theta \\ &= 36\int_0^{2\pi} 1 - 2\cos\theta + \cos^2\theta \, d\theta \\ &= 36\int_0^{2\pi} \frac{3}{2} - 2\cos\theta + \frac{1}{2}\cos(2\theta) \, d\theta \\ &= 36\left(\frac{3}{2}\theta - 2\sin\theta + \frac{1}{4}\sin(2\theta)\right)\Bigg|_0^{2\pi} \\ &= 108\pi \end{aligned}$$

The parametric curve (without the limits) we used in the previous example is called a **cycloid**. In its general form the cycloid is,

$$x = r(\theta - \sin\theta) \qquad y = r(1 - \cos\theta)$$

The cycloid represents the following situation. Consider a wheel of radius *r*. Let the point where the wheel touches the ground initially be called *P*. Then start rolling the wheel to the right. As the wheel rolls to the right trace out the path of the point *P*. The path that the point *P* traces out is called a cycloid and is given by the equations above. In these equations we can think of θ as the angle through which the point *P* has rotated.

Here is a cycloid sketched out with the wheel shown at various places. The blue dot is the point *P* on the wheel that we're using to trace out the curve.



From this sketch we can see that one arch of the cycloid is traced out in the range $0 \le \theta \le 2\pi$. This makes sense when you consider that the point P will be back on the ground after it has rotated through an angle of 2π.

Arc Length with Parametric Equations

In the previous two sections we've looked at a couple of Calculus I topics in terms of parametric equations. We now need to look at a couple of Calculus II topics in terms of parametric equations.

In this section we will look at the arc length of the parametric curve given by,

$$x = f(t) \qquad\qquad y = g(t) \qquad\qquad \alpha \le t \le \beta$$

We will also be assuming that the curve is traced out exactly once as t increases from α to β. We will also need to assume that the curve is traced out from left to right as t increases. This is equivalent to saying,

$$\frac{dx}{dt} \ge 0 \qquad\qquad \text{for } \alpha \le t \le \beta$$

So, let's start out the derivation by recalling the arc length formula as we first derived it in the arc length section of the Applications of Integrals chapter.

$$L = \int ds$$

where,

$$ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx \qquad \text{if } y = f(x),\ a \le x \le b$$

$$ds = \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\,dy \qquad \text{if } x = h(y),\ c \le y \le d$$

We will use the first ds above because we have a nice formula for the derivative in terms of the parametric equations (see the Tangents with Parametric Equations section). To use this we'll also need to know that,

$$dx = f'(t)\,dt = \frac{dx}{dt}\,dt$$

The arc length formula then becomes,

$$L = \int_{\alpha}^{\beta} \sqrt{1 + \left(\frac{\frac{dy}{dt}}{\frac{dx}{dt}}\right)^2}\ \frac{dx}{dt}\,dt = \int_{\alpha}^{\beta} \sqrt{1 + \frac{\left(\frac{dy}{dt}\right)^2}{\left(\frac{dx}{dt}\right)^2}}\ \frac{dx}{dt}\,dt$$

This is a particularly unpleasant formula. However, if we factor out the denominator from the square root we arrive at,

$$L = \int_{\alpha}^{\beta} \frac{1}{\left|\frac{dx}{dt}\right|} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \; \frac{dx}{dt} dt$$

Now, making use of our assumption that the curve is being traced out from left to right we can drop the absolute value bars on the derivative which will allow us to cancel the two derivatives that are outside the square root and this gives,

Arc Length for Parametric Equations

$$L = \int_{\alpha}^{\beta} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \; dt$$

Notice that we could have used the second formula for *ds* above if we had assumed instead that

$$\frac{dy}{dt} \geq 0 \qquad \text{for} \quad \alpha \leq t \leq \beta$$

If we had gone this route in the derivation we would have gotten the same formula.

Let's take a look at an example.

Example 1 Determine the length of the parametric curve given by the following parametric equations.

$$x = 3\sin(t) \qquad\qquad y = 3\cos(t) \qquad\qquad 0 \leq t \leq 2\pi$$

Solution

We know that this is a circle of radius 3 centered at the origin from our prior discussion about graphing parametric curves. We also know from this discussion that it will be traced out exactly once in this range.

So, we can use the formula we derived above. We'll first need the following,

$$\frac{dx}{dt} = 3\cos(t) \qquad\qquad \frac{dy}{dt} = -3\sin(t)$$

The length is then,

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{9\sin^2(t) + 9\cos^2(t)} \; dt \\ &= \int_0^{2\pi} 3\sqrt{\sin^2(t) + \cos^2(t)} \; dt \\ &= 3\int_0^{2\pi} dt \\ &= 6\pi \end{aligned}$$

Since this is a circle we could have just used the fact that the length of the circle is just the circumference of the circle. This is a nice way, in this case, to verify our result.

Let's take a look at one possible consequence if a curve is traced out more than once and we try to find the length of the curve without taking this into account.

Example 2 Use the arc length formula for the following parametric equations.

$$x = 3\sin(3t) \qquad y = 3\cos(3t) \qquad 0 \le t \le 2\pi$$

Solution

Notice that this is the identical circle that we had in the previous example and so the length is still 6π. However, for the range given we know it will trace out the curve three times instead once as required for the formula. Despite that restriction let's use the formula anyway and see what happens.

In this case the derivatives are,

$$\frac{dx}{dt} = 9\cos(3t) \qquad \frac{dy}{dt} = -9\sin(3t)$$

and the length formula gives,

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{81\sin^2(3t) + 81\cos^2(3t)}\, dt \\ &= \int_0^{2\pi} 9\, dt \\ &= 18\pi \end{aligned}$$

The answer we got form the arc length formula in this example was 3 times the actual length. Recalling that we also determined that this circle would trace out three times in the range given, the answer should make some sense.

If we had wanted to determine the length of the circle for this set of parametric equations we would need to determine a range of t for which this circle is traced out exactly once. This is, $0 \le t \le \frac{2\pi}{3}$. Using this range of t we get the following for the length.

$$\begin{aligned} L &= \int_0^{\frac{2\pi}{3}} \sqrt{81\sin^2(3t) + 81\cos^2(3t)}\, dt \\ &= \int_0^{\frac{2\pi}{3}} 9\, dt \\ &= 6\pi \end{aligned}$$

which is the correct answer.

Be careful to not make the assumption that this is always what will happen if the curve is traced out more than once. Just because the curve traces out n times does not mean that the arc length formula will give us n times the actual length of the curve!

Before moving on to the next section let's notice that we can put the arc length formula derived in this section into the same form that we had when we first looked at arc length. The only difference is that we will add in a definition for ds when we have parametric equations.

The arc length formula can be summarized as,

$$L = \int ds$$

where,

$$ds = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \qquad \text{if } y = f(x),\ a \le x \le b$$

$$ds = \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy \qquad \text{if } x = h(y),\ c \le y \le d$$

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \qquad \text{if } x = f(t), y = g(t),\ \alpha \le t \le \beta$$

Surface Area with Parametric Equations

In this final section of looking at calculus applications with parametric equations we will take a look at determining the surface area of a region obtained by rotating a parametric curve about the *x* or *y*-axis.

We will rotate the parametric curve given by,

$$x = f(t) \qquad\qquad y = g(t) \qquad\qquad \alpha \le t \le \beta$$

about the *x* or *y*-axis. We are going to assume that the curve is traced out exactly once as *t* increases from α to β. At this point there actually isn't all that much to do. We know that the surface area can be found by using one of the following two formulas depending on the axis of rotation (recall the Surface Area section of the Applications of Integrals chapter).

$$S = \int 2\pi y \, ds \qquad\qquad \text{rotation about } x-\text{axis}$$

$$S = \int 2\pi x \, ds \qquad\qquad \text{rotation about } y-\text{axis}$$

All that we need is a formula for *ds* to use and from the previous section we have,

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \qquad \text{if } x = f(t), y = g(t), \ \alpha \le t \le \beta$$

which is exactly what we need.

We will need to be careful with the *x* or *y* that is in the original surface area formula. Back when we first looked at surface area we saw that sometimes we had to substitute for the variable in the integral and at other times we didn't. This was dependent upon the *ds* that we used. In this case however, we will always have to substitute for the variable. The *ds* that we use for parametric equations introduces a *dt* into the integral and that means that everything needs to be in terms of *t*. Therefore, we will need to substitute the appropriate parametric equation for *x* or *y* depending on the axis of rotation.

Let's take a quick look at an example.

Example 1 Determine the surface area of the solid obtained by rotating the following parametric curve about the *x*-axis.

$$x = \cos^3\theta \qquad y = \sin^3\theta \qquad 0 \le \theta \le \frac{\pi}{2}$$

Solution

We'll first need the derivatives of the parametric equations.

$$\frac{dx}{dt} = -3\cos^2\theta\sin\theta \qquad\qquad \frac{dy}{dt} = 3\sin^2\theta\cos\theta$$

Before plugging into the surface area formula let's get the *ds* out of the way.

$$
\begin{aligned}
ds &= \sqrt{9\cos^4\theta\sin^2\theta + 9\sin^4\theta\cos^2\theta}\, d\theta \\
&= 3\left|\cos\theta\sin\theta\right|\sqrt{\cos^2\theta + \sin^2\theta} \\
&= 3\cos\theta\sin\theta
\end{aligned}
$$

Notice that we could drop the absolute value bars since both sine and cosine are positive in this range of θ given.

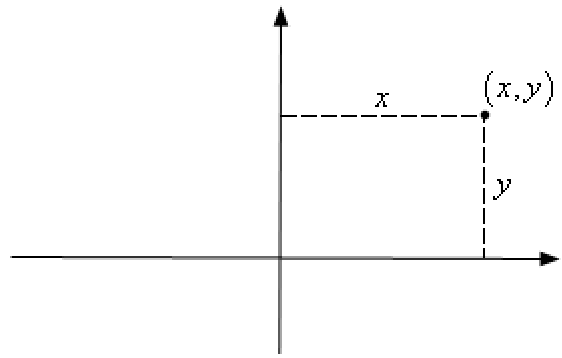
Now let's get the surface area and don't forget to also plug in for the y.

$$
\begin{aligned}
S &= \int 2\pi y\, ds \\
&= 2\pi\int_0^{\frac{\pi}{2}} \sin^3\theta\left(3\cos\theta\sin\theta\right)\, d\theta \\
&= 6\pi\int_0^{\frac{\pi}{2}} \sin^4\theta\cos\theta\, d\theta \qquad\qquad u = \sin\theta \\
&= 6\pi\int_0^1 u^4\, du \\
&= \frac{6\pi}{5}
\end{aligned}
$$

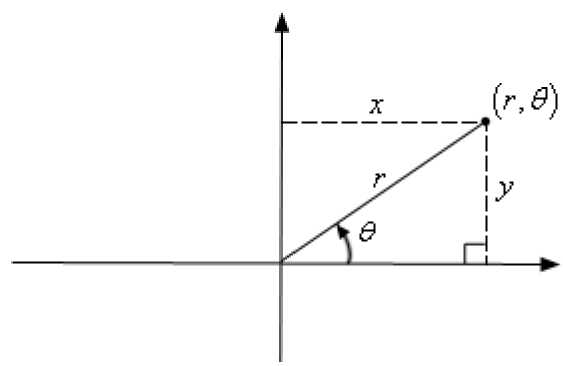
Polar Coordinates

Up to this point we've dealt exclusively with the Cartesian (or Rectangular, or *x*-*y*) coordinate system. However, as we will see, this is not always the easiest coordinate system to work in. So, in this section we will start looking at the polar coordinate system.

Coordinate systems are really nothing more than a way to define a point in space. For instance in the Cartesian coordinate system at point is given the coordinates (x,y) and we use this to define the point by starting at the origin and then moving x units horizontally followed by y units vertically. This is shown in the sketch below.

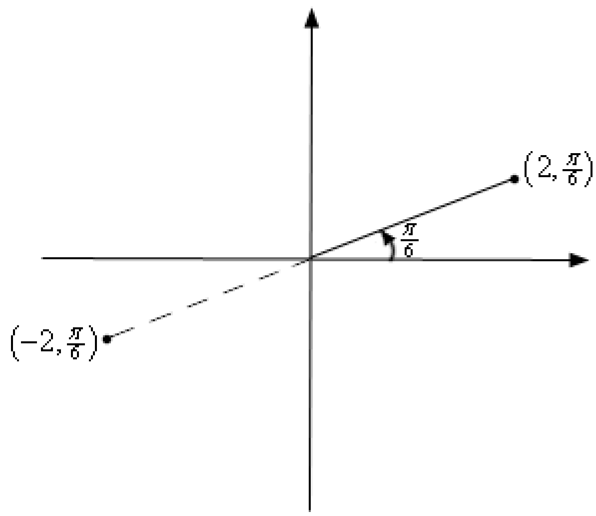


This is not, however, the only way to define a point in two dimensional space. Instead of moving vertically and horizontally from the origin to get to the point we could instead go straight out of the origin until we hit the point and then determine the angle this line makes with the positive *x*-axis. We could then use the distance of the point from the origin and the amount we needed to rotate from the positive *x*-axis as the coordinates of the point. This is shown in the sketch below.



Coordinates in this form are called **polar coordinates**.

The above discussion may lead one to think that r must be a positive number. However, we also allow r to be negative. Below is a sketch of the two points $\left(2,\frac{\pi}{6}\right)$ and $\left(-2,\frac{\pi}{6}\right)$.

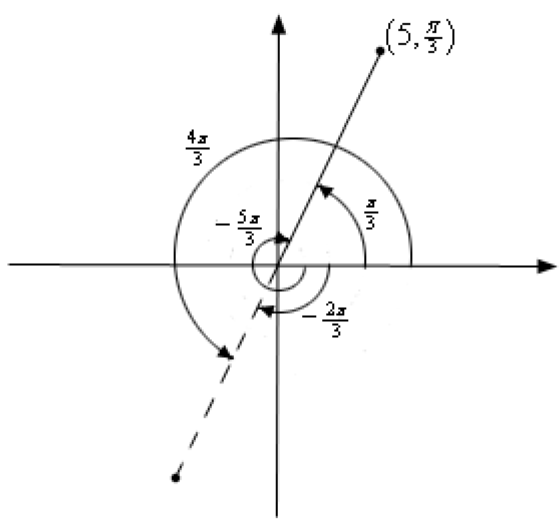


From this sketch we can see that if r is positive the point will be in the same quadrant as θ. On the other hand if r is negative the point will end up in the quadrant exactly opposite θ. Notice as well that the coordinates $\left(-2,\frac{\pi}{6}\right)$ describe the same point as the coordinates $\left(2,\frac{7\pi}{6}\right)$ do. The coordinates $\left(2,\frac{7\pi}{6}\right)$ tells us to rotate an angle of $\frac{7\pi}{6}$ from the positive x-axis, this would put us on the dashed line in the sketch above, and then move out a distance of 2.

This leads to an important difference between Cartesian coordinates and polar coordinates. In Cartesian coordinates there is exactly one set of coordinates for any given point. With polar coordinates this isn't true. In polar coordinates there is literally an infinite number of coordinates for a given point. For instance, the following four points are all coordinates for the same point.

$$\left(5,\frac{\pi}{3}\right)=\left(5,-\frac{5\pi}{3}\right)=\left(-5,\frac{4\pi}{3}\right)=\left(-5,-\frac{2\pi}{3}\right)$$

Here is a sketch of the angles used in these four sets of coordinates.



In the second coordinate pair we rotated in a clock-wise direction to get to the point. We shouldn't forget about rotating in the clock-wise direction. Sometimes it's what we have to do.

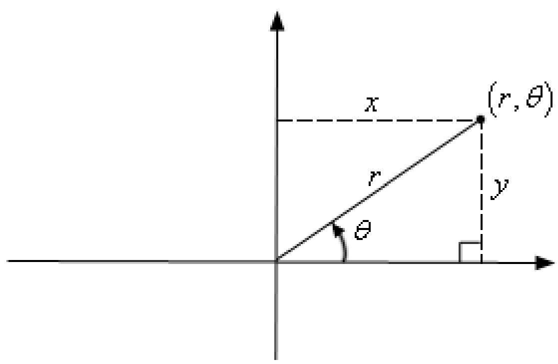
The last two coordinate pairs use the fact that if we end up in the opposite quadrant from the point we can use a negative *r* to get back to the point and of course there is both a counter clock-wise and a clock-wise rotation to get to the angle.

These four points only represent the coordinates of the point without rotating around the system more than once. If we allow the angle to make as many complete rotations about the axis system as we want then there are an infinite number of coordinates for the same point. In fact the point (r,θ) can be represented by any of the following coordinate pairs.

$$(r,\theta+2\pi n) \qquad\qquad (-r,\theta+(2n+1)\pi), \qquad \text{where } n \text{ is any integer.}$$

Next we should talk about the origin of the coordinate system. In polar coordinates the origin is often called the **pole**. Because we aren't actually moving away from the origin/pole we know that $r=0$. However, we can still rotate around the system by any angle we want and so the coordinates of the origin/pole are $(0,\theta)$.

Now that we've got a grasp on polar coordinates we need to think about converting between the two coordinate systems. Well start out with the following sketch reminding us how both coordinate systems work.



Note that we've got a right triangle above and with that we can get the following equations that will convert polar coordinates into Cartesian coordinates.

Polar to Cartesian Conversion Formulas

$$x = r\cos\theta \qquad\qquad y = r\sin\theta$$

Converting from Cartesian is almost as easy. Let's first notice the following.

$$\begin{aligned} x^2+y^2 &= (r\cos\theta)^2+(r\sin\theta)^2 \\ &= r^2\cos^2\theta+r^2\sin^2\theta \\ &= r^2\left(\cos^2\theta+\sin^2\theta\right)=r^2 \end{aligned}$$

This is a very useful formula that we should remember, however we are after an equation for *r* so let's take the square root of both sides. This gives,

$$r=\sqrt{x^2+y^2}$$

Note that technically we should have a plus or minus in front of the root since we know that r can be either positive or negative. We will run with the convention of positive r here.

Getting an equation for θ is almost as simple. We'll start with,

$$\frac{y}{x} = \frac{r\sin\theta}{r\cos\theta} = \tan\theta$$

Taking the inverse tangent of both sides gives,

$$\theta = \tan^{-1}\left(\frac{y}{x}\right)$$

We will need to be careful with this because inverse tangents only return values in the range $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$. Recall that there is a second possible angle and that the second angle is given by $\theta + \pi$.

Summarizing then gives the following formulas for converting from Cartesian coordinates to polar coordinates.

Cartesian to Polar Conversion Formulas

$$r^2 = x^2 + y^2 \qquad\qquad r = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1}\left(\frac{y}{x}\right)$$

Let's work a quick example.

Example 1 Convert each of the following points into the given coordinate system.

(a) $\left(-4, \frac{2\pi}{3}\right)$ into Cartesian coordinates. [Solution]

(b) (-1,-1) into polar coordinates. [Solution]

Solution

(a) Convert $\left(-4, \frac{2\pi}{3}\right)$ **into Cartesian coordinates.**

This conversion is easy enough. All we need to do is plug the points into the formulas.

$$x = -4\cos\left(\frac{2\pi}{3}\right) = -4\left(-\frac{1}{2}\right) = 2$$

$$y = -4\sin\left(\frac{2\pi}{3}\right) = -4\left(\frac{\sqrt{3}}{2}\right) = -2\sqrt{3}$$

So, in Cartesian coordinates this point is $\left(2, -2\sqrt{3}\right)$.

[Return to Problems]

(b) Convert (-1,-1) **into polar coordinates.**

Let's first get *r*.

$$r = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}$$

Now, let's get θ.

$$\theta = \tan^{-1}\left(\frac{-1}{-1}\right) = \tan^{-1}(1) = \frac{\pi}{4}$$

This is not the correct angle however. This value of θ is in the first quadrant and the point we've been given is in the third quadrant. As noted above we can get the correct angle by adding π onto this. Therefore, the actual angle is,

$$\theta = \frac{\pi}{4} + \pi = \frac{5\pi}{4}$$

So, in polar coordinates the point is $\left(\sqrt{2}, \frac{5\pi}{4}\right)$. Note as well that we could have used the first θ that we got by using a negative *r*. In this case the point could also be written in polar coordinates as $\left(-\sqrt{2}, \frac{\pi}{4}\right)$.

[Return to Problems]

We can also use the above formulas to convert equations from one coordinate system to the other.

Example 2 Convert each of the following into an equation in the given coordinate system.

(a) Convert $2x - 5x^3 = 1 + xy$ into polar coordinates. [Solution]

(b) Convert $r = -8\cos\theta$ into Cartesian coordinates. [Solution]

Solution

(a) Convert $2x - 5x^3 = 1 + xy$ **into polar coordinates.**

In this case there really isn't much to do other than plugging in the formulas for *x* and *y* (*i.e.* the Cartesian coordinates) in terms of *r* and θ (*i.e.* the polar coordinates).

$$2(r\cos\theta) - 5(r\cos\theta)^3 = 1 + (r\cos\theta)(r\sin\theta)$$

$$2r\cos\theta - 5r^3\cos^3\theta = 1 + r^2\cos\theta\sin\theta$$

[Return to Problems]

(b) Convert $r = -8\cos\theta$ **into Cartesian coordinates.**

This one is a little trickier, but not by much. First notice that we could substitute straight for the *r*. However, there is no straight substitution for the cosine that will give us only Cartesian coordinates. If we had an *r* on the right along with the cosine then we could do a direct substitution. So, if an *r* on the right side would be convenient let's put one there, just don't forget to put one on the left side as well.

$$r^2 = -8r\cos\theta$$

We can now make some substitutions that will convert this into Cartesian coordinates.

$$x^2 + y^2 = -8x$$

[Return to Problems]

Before moving on to the next subject let's do a little more work on the second part of the previous example.

The equation given in the second part is actually a fairly well known graph; it just isn't in a form that most people will quickly recognize. To identify it let's take the Cartesian coordinate equation and do a little rearranging.

$$x^2 + 8x + y^2 = 0$$

Now, complete the square on the *x* portion of the equation.

$$\begin{aligned} x^2 + 8x + 16 + y^2 &= 16 \\ (x+4)^2 + y^2 &= 16 \end{aligned}$$

So, this was a circle of radius 4 and center (-4,0).

This leads us into the final topic of this section.

Common Polar Coordinate Graphs

Let's identify a few of the more common graphs in polar coordinates. We'll also take a look at a couple of special polar graphs.

Lines

Some lines have fairly simple equations in polar coordinates.

1. $\theta = \beta$.

We can see that this is a line by converting to Cartesian coordinates as follows

$$\begin{aligned} \theta &= \beta \\ \tan^{-1}\left(\frac{y}{x}\right) &= \beta \\ \frac{y}{x} &= \tan\beta \\ y &= (\tan\beta)\, x \end{aligned}$$

This is a line that goes through the origin and makes an angle of β with the positive *x*-axis. Or, in other words it is a line through the origin with slope of $\tan\beta$.

2. $r\cos\theta = a$

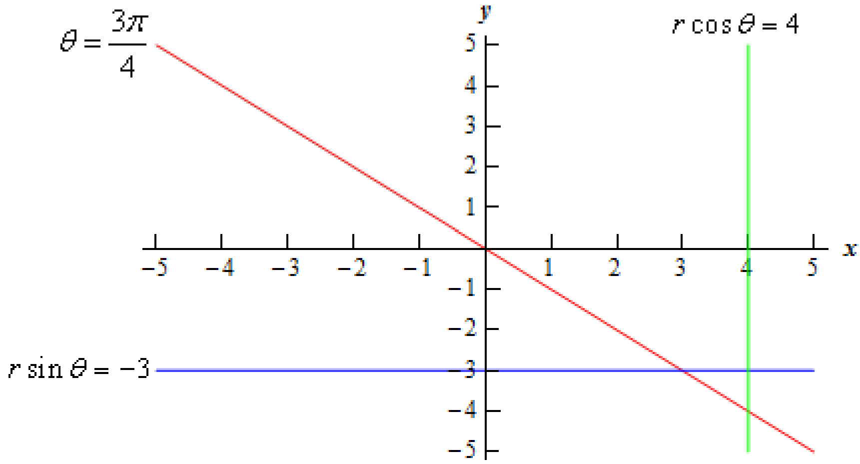
This is easy enough to convert to Cartesian coordinates to $x = a$. So, this is a vertical line.

3. $r\sin\theta = b$

Likewise, this converts to $y = b$ and so is a horizontal line.

Example 3 Graph $\theta = \frac{3\pi}{4}$, $r\cos\theta = 4$ and $r\sin\theta = -3$ on the same axis system.

Solution
There really isn't too much to this one other than doing the graph so here it is.



Circles
Let's take a look at the equations of circles in polar coordinates.

1. $r = a$.
This equation is saying that no matter what angle we've got the distance from the origin must be *a*. If you think about it that is exactly the definition of a circle of radius *a* centered at the origin.

So, this is a circle of radius *a* centered at the origin. This is also one of the reasons why we might want to work in polar coordinates. The equation of a circle centered at the origin has a very nice equation, unlike the corresponding equation in Cartesian coordinates.

2. $r = 2a\cos\theta$.
We looked at a specific example of one of these when we were converting equations to Cartesian coordinates.

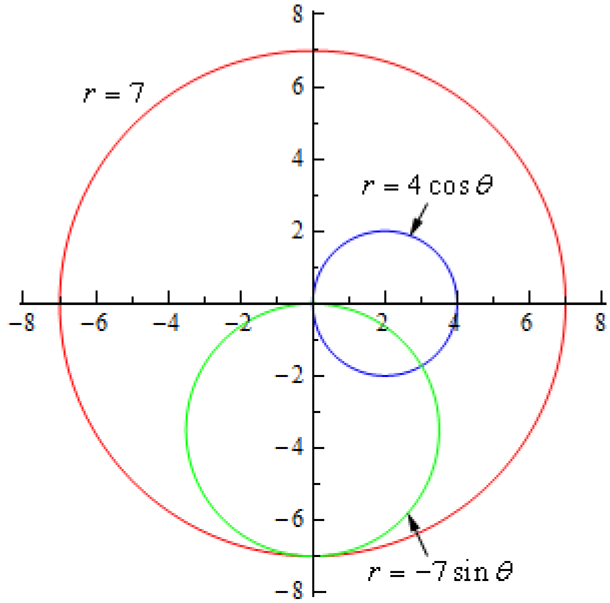
This is a circle of radius $|a|$ and center $(a,0)$. Note that *a* might be negative (as it was in our example above) and so the absolute value bars are required on the radius. They should not be used however on the center.

3. $r = 2b\sin\theta$.
This is similar to the previous one. It is a circle of radius $|b|$ and center $(0,b)$.

4. $r = 2a\cos\theta + 2b\sin\theta$.
This is a combination of the previous two and by completing the square twice it can be shown that this is a circle of radius $\sqrt{a^2+b^2}$ and center (a,b). In other words, this is the general equation of a circle that isn't centered at the origin.

Example 4 Graph $r = 7$, $r = 4\cos\theta$, and $r = -7\sin\theta$ on the same axis system.

Solution
The first one is a circle of radius 7 centered at the origin. The second is a circle of radius 2 centered at (2,0). The third is a circle of radius $\frac{7}{2}$ centered at $\left(0,-\frac{7}{2}\right)$. Here is the graph of the three equations.



Note that it takes a range of $0 \le \theta \le 2\pi$ for a complete graph of $r = a$ and it only takes a range of $0 \le \theta \le \pi$ to graph the other circles given here.

Cardioids and Limacons
These can be broken up into the following three cases.

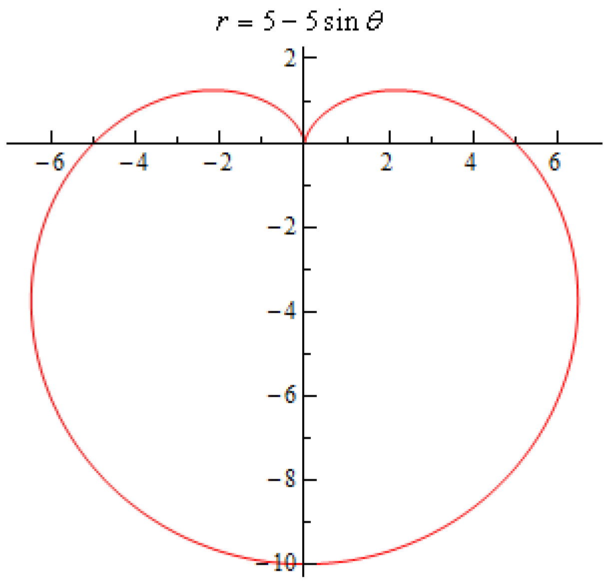
1. Cardioids : $r = a \pm a\cos\theta$ and $r = a \pm a\sin\theta$.
   These have a graph that is vaguely heart shaped and always contain the origin.

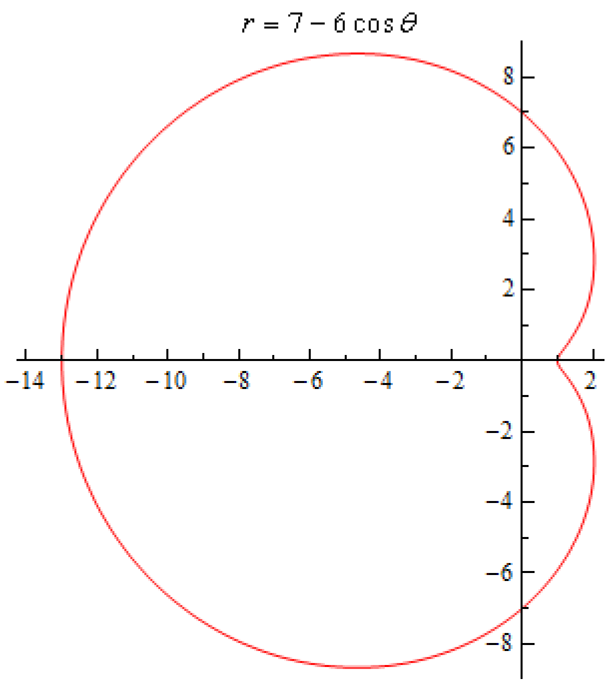
2. Limacons with an inner loop : $r = a \pm b\cos\theta$ and $r = a \pm b\sin\theta$ with $a < b$.
   These will have an inner loop and will always contain the origin.

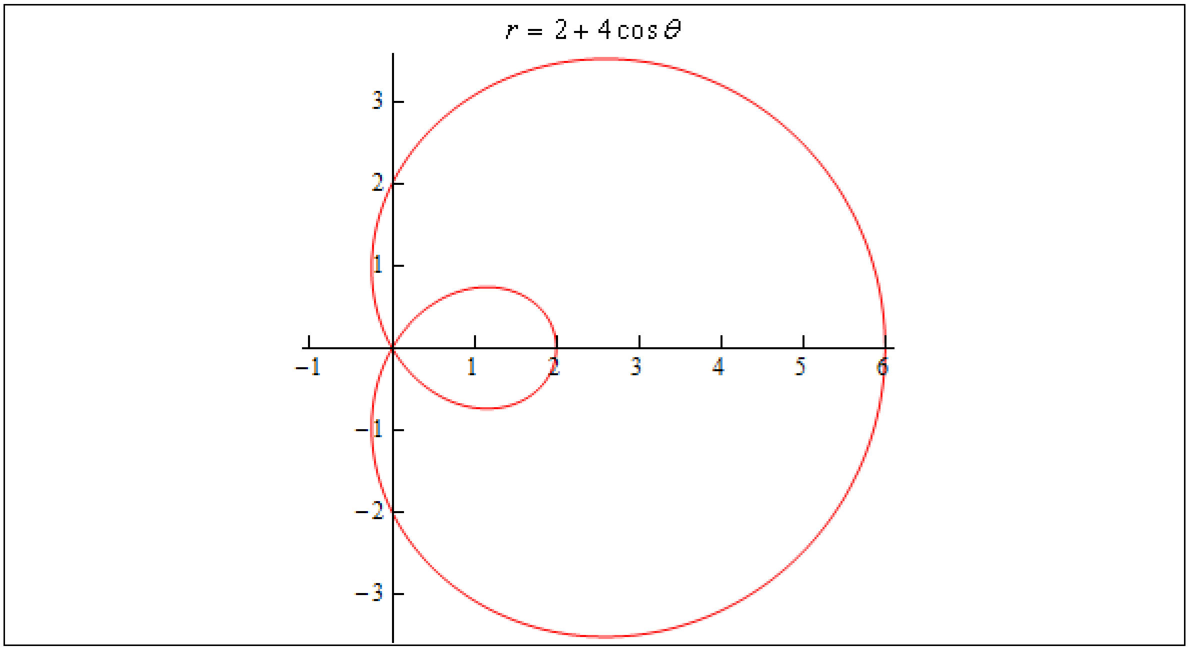
3. Limacons without an inner loop : $r = a \pm b\cos\theta$ and $r = a \pm b\sin\theta$ with $a > b$.
   These do not have an inner loop and do not contain the origin.

Example 5 Graph $r = 5 - 5\sin\theta$, $r = 7 - 6\cos\theta$, and $r = 2 + 4\cos\theta$.
Solution
These will all graph out once in the range $0 \le \theta \le 2\pi$. Here is a table of values for each followed by graphs of each.

θ	$r = 5 - 5\sin\theta$	$r = 7 - 6\cos\theta$	$r = 2 + 4\cos\theta$
0	5	1	6
$\frac{\pi}{2}$	0	7	2
π	5	13	-2
$\frac{3\pi}{2}$	10	7	2
2π	5	1	6







There is one final thing that we need to do in this section. In the third graph in the previous example we had an inner loop. We will, on occasion, need to know the value of θ for which the graph will pass through the origin. To find these all we need to do is set the equation equal to zero and solve as follows,

$$0 = 2 + 4\cos\theta \qquad \Rightarrow \qquad \cos\theta = -\frac{1}{2} \qquad \Rightarrow \qquad \theta = \frac{2\pi}{3}, \frac{4\pi}{3}$$

Tangents with Polar Coordinates

We now need to discuss some calculus topics in terms of polar coordinates.

We will start with finding tangent lines to polar curves. In this case we are going to assume that the equation is in the form $r = f\left(\theta\right)$. With the equation in this form we can actually use the equation for the derivative $\frac{dy}{dx}$ we derived when we looked at tangent lines with parametric equations. To do this however requires us to come up with a set of parametric equations to represent the curve. This is actually pretty easy to do.

From our work in the previous section we have the following set of conversion equations for going from polar coordinates to Cartesian coordinates.

$$x = r\cos\theta \qquad\qquad y = r\sin\theta$$

Now, we'll use the fact that we're assuming that the equation is in the form $r = f\left(\theta\right)$. Substituting this into these equations gives the following set of parametric equations (with θ as the parameter) for the curve.

$$x = f\left(\theta\right)\cos\theta \qquad\qquad y = f\left(\theta\right)\sin\theta$$

Now, we will need the following derivatives.

$$\begin{aligned}\frac{dx}{d\theta} &= f'\left(\theta\right)\cos\theta - f\left(\theta\right)\sin\theta \\ &= \frac{dr}{d\theta}\cos\theta - r\sin\theta\end{aligned} \qquad\qquad \begin{aligned}\frac{dy}{d\theta} &= f'\left(\theta\right)\sin\theta + f\left(\theta\right)\cos\theta \\ &= \frac{dr}{d\theta}\sin\theta + r\cos\theta\end{aligned}$$

The derivative $\frac{dy}{dx}$ is then,

Derivative with Polar Coordinates

$$\frac{dy}{dx} = \frac{\frac{dr}{d\theta}\sin\theta + r\cos\theta}{\frac{dr}{d\theta}\cos\theta - r\sin\theta}$$

Note that rather than trying to remember this formula it would probably be easier to remember how we derived it and just remember the formula for parametric equations.

Let's work a quick example with this.

Example 1 Determine the equation of the tangent line to $r = 3 + 8\sin\theta$ at $\theta = \dfrac{\pi}{6}$.

Solution
We'll first need the following derivative.

$$\frac{dr}{d\theta} = 8\cos\theta$$

The formula for the derivative $\dfrac{dy}{dx}$ becomes,

$$\frac{dy}{dx} = \frac{8\cos\theta\sin\theta + \left(3 + 8\sin\theta\right)\cos\theta}{8\cos^2\theta - \left(3 + 8\sin\theta\right)\sin\theta} = \frac{16\cos\theta\sin\theta + 3\cos\theta}{8\cos^2\theta - 3\sin\theta - 8\sin^2\theta}$$

The slope of the tangent line is,

$$m = \left.\frac{dy}{dx}\right|_{\theta = \frac{\pi}{6}} = \frac{4\sqrt{3} + \frac{3\sqrt{3}}{2}}{4 - \frac{3}{2}} = \frac{11\sqrt{3}}{5}$$

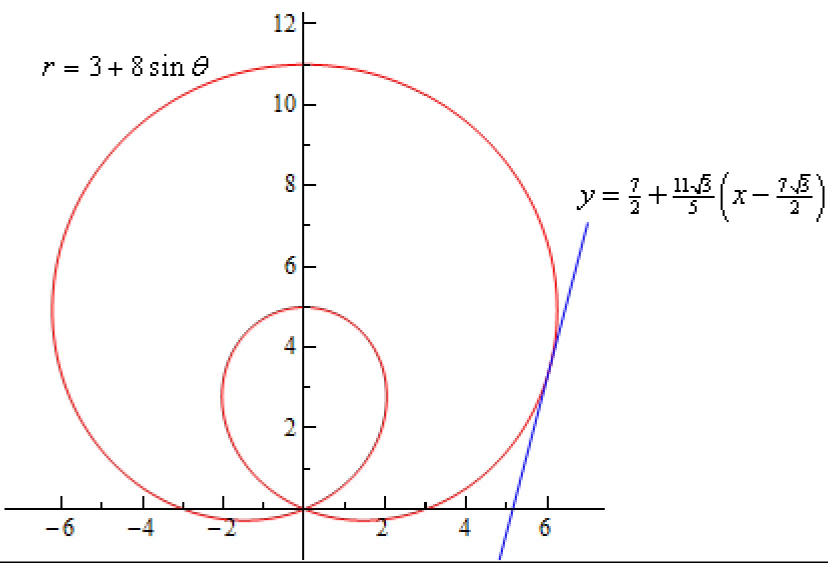
Now, at $\theta = \frac{\pi}{6}$ we have $r = 7$. We'll need to get the corresponding *x*-*y* coordinates so we can get the tangent line.

$$x = 7\cos\left(\frac{\pi}{6}\right) = \frac{7\sqrt{3}}{2} \qquad\qquad y = 7\sin\left(\frac{\pi}{6}\right) = \frac{7}{2}$$

The tangent line is then,

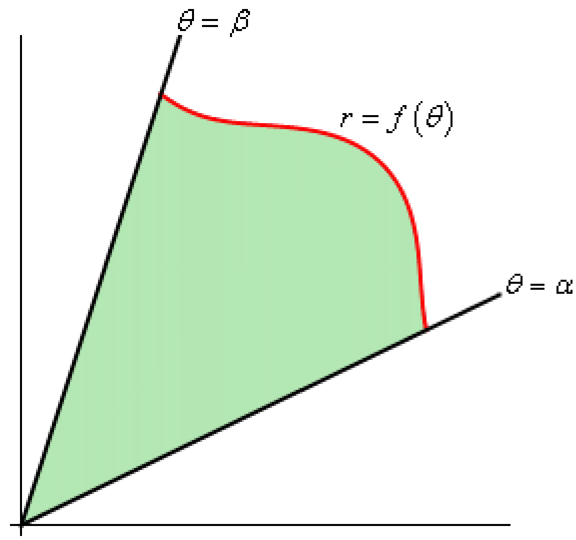
$$y = \frac{7}{2} + \frac{11\sqrt{3}}{5}\left(x - \frac{7\sqrt{3}}{2}\right)$$

For the sake of completeness here is a graph of the curve and the tangent line.



Area with Polar Coordinates

In this section we are going to look at areas enclosed by polar curves. Note as well that we said "enclosed by" instead of "under" as we typically have in these problems. These problems work a little differently in polar coordinates. Here is a sketch of what the area that we'll be finding in this section looks like.



We'll be looking for the shaded area in the sketch above. The formula for finding this area is,

$$A = \int_{\alpha}^{\beta} \frac{1}{2} r^2 \, d\theta$$

Notice that we use *r* in the integral instead of $f(\theta)$ so make sure and substitute accordingly when doing the integral.

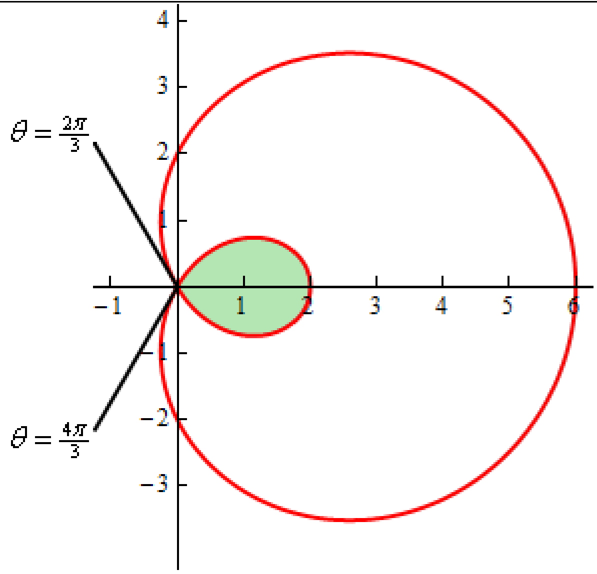
Let's take a look at an example.

Example 1 Determine the area of the inner loop of $r = 2 + 4\cos\theta$.

Solution
We graphed this function back when we first started looking at polar coordinates. For this problem we'll also need to know the values of θ where the curve goes through the origin. We can get these by setting the equation equal to zero and solving.

$$0 = 2 + 4\cos\theta$$
$$\cos\theta = -\frac{1}{2} \qquad \Rightarrow \qquad \theta = \frac{2\pi}{3}, \frac{4\pi}{3}$$

Here is the sketch of this curve with the inner loop shaded in.



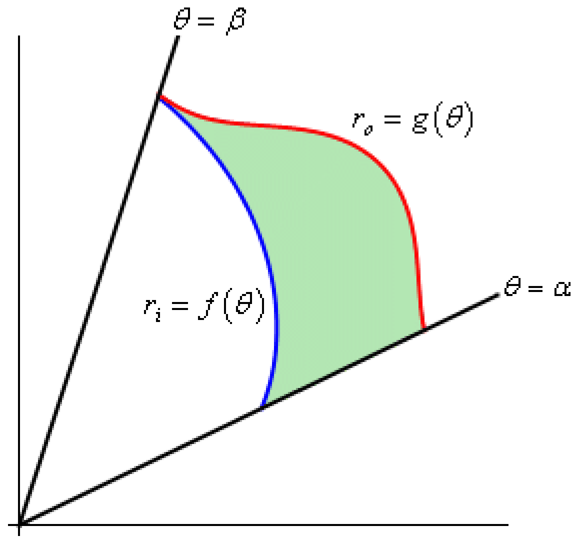
Can you see why we needed to know the values of θ where the curve goes through the origin? These points define where the inner loop starts and ends and hence are also the limits of integration in the formula.

So, the area is then,

$$\begin{aligned} A &= \int_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} \frac{1}{2}\left(2+4\cos\theta\right)^2 d\theta \\ &= \int_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} \frac{1}{2}\left(4+16\cos\theta+16\cos^2\theta\right) d\theta \\ &= \int_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} 2+8\cos\theta+4\left(1+\cos\left(2\theta\right)\right) d\theta \\ &= \int_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} 6+8\cos\theta+4\cos\left(2\theta\right) d\theta \\ &= \left(6\theta+8\sin\theta+2\sin\left(2\theta\right)\right)\Big|_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} \\ &= 4\pi-6\sqrt{3} = 2.174 \end{aligned}$$

You did follow the work done in this integral didn't you? You'll run into quite a few integrals of trig functions in this section so if you need to you should go back to the Integrals Involving Trig Functions sections and do a quick review.

So, that's how we determine areas that are enclosed by a single curve, but what about situations like the following sketch were we want to find the area between two curves.



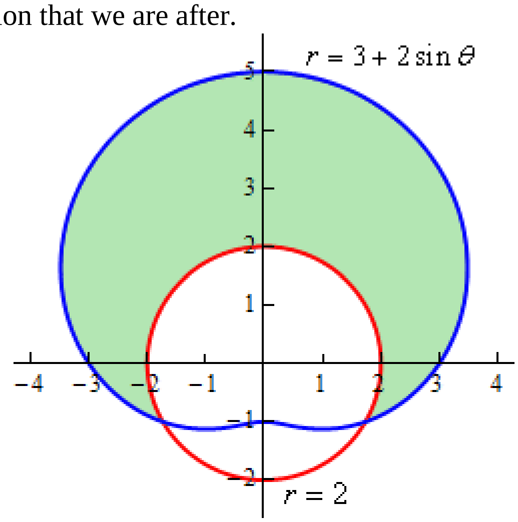
In this case we can use the above formula to find the area enclosed by both and then the actual area is the difference between the two. The formula for this is,

$$A = \int_{\alpha}^{\beta} \frac{1}{2}\left(r_o^2 - r_i^2\right) d\theta$$

Let's take a look at an example of this.

Example 2 Determine the area that lies inside $r = 3 + 2\sin\theta$ and outside $r = 2$.

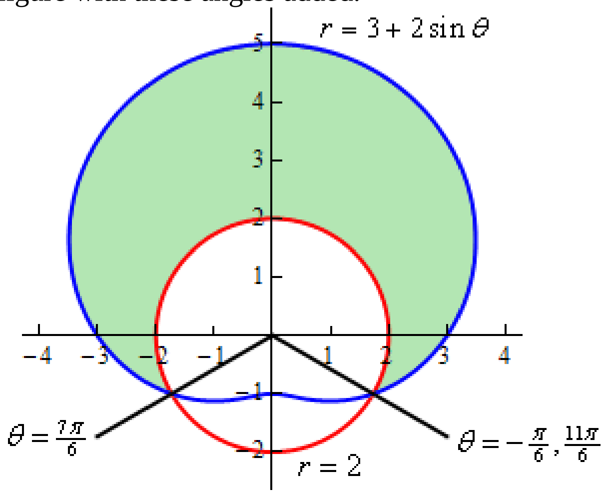
Solution
Here is a sketch of the region that we are after.



To determine this area we'll need to know the values of θ for which the two curves intersect. We can determine these points by setting the two equations and solving.

$$3+2\sin\theta = 2$$
$$\sin\theta = -\frac{1}{2} \qquad \Rightarrow \qquad \theta = \frac{7\pi}{6}, \frac{11\pi}{6}$$

Here is a sketch of the figure with these angles added.



Note as well here that we also acknowledged that another representation for the angle $\frac{11\pi}{6}$ is $-\frac{\pi}{6}$. This is important for this problem. In order to use the formula above the area must be enclosed as we increase from the smaller to larger angle. So, if we use $\frac{7\pi}{6}$ to $\frac{11\pi}{6}$ we will not enclose the shaded area, instead we will enclose the bottom most of the three regions. However, if we use the angles $-\frac{\pi}{6}$ to $\frac{7\pi}{6}$ we will enclose the area that we're after.

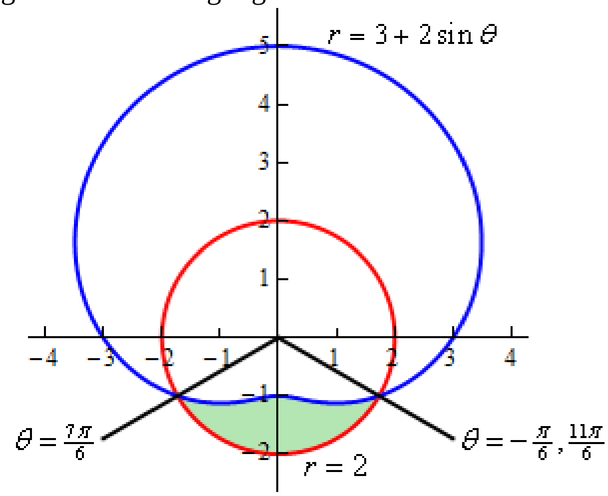
So, the area is then,

$$\begin{aligned} A &= \int_{-\frac{\pi}{6}}^{\frac{7\pi}{6}} \frac{1}{2}\left(\left(3+2\sin\theta\right)^2 - \left(2\right)^2\right)d\theta \\ &= \int_{-\frac{\pi}{6}}^{\frac{7\pi}{6}} \frac{1}{2}\left(5+12\sin\theta+4\sin^2\theta\right)d\theta \\ &= \int_{-\frac{\pi}{6}}^{\frac{7\pi}{6}} \frac{1}{2}\left(7+12\sin\theta-2\cos\left(2\theta\right)\right)d\theta \\ &= \frac{1}{2}\left(7\theta-12\cos\theta-\sin\left(2\theta\right)\right)\Big|_{-\frac{\pi}{6}}^{\frac{7\pi}{6}} \\ &= \frac{11\sqrt{3}}{2}+\frac{14\pi}{3} = 24.187 \end{aligned}$$

Let's work a slight modification of the previous example.

Example 3 Determine the area of the region outside $r = 3 + 2\sin\theta$ and inside $r = 2$.

Solution
This time we're looking for the following region.



So, this is the region that we get by using the limits $\frac{7\pi}{6}$ to $\frac{11\pi}{6}$. The area for this region is,

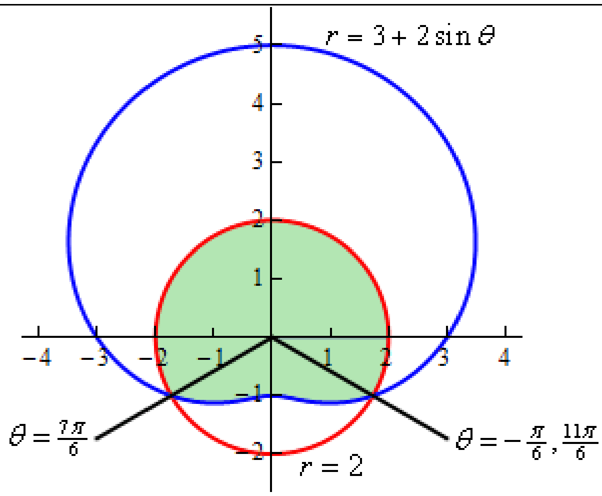
$$\begin{aligned} A &= \int_{\frac{7\pi}{6}}^{\frac{11\pi}{6}} \frac{1}{2}\left((2)^2 - (3 + 2\sin\theta)^2 \right) d\theta \\ &= \int_{\frac{7\pi}{6}}^{\frac{11\pi}{6}} \frac{1}{2}\left(-5 - 12\sin\theta - 4\sin^2\theta \right) d\theta \\ &= \int_{\frac{7\pi}{6}}^{\frac{11\pi}{6}} \frac{1}{2}\left(-7 - 12\sin\theta + 2\cos(2\theta) \right) d\theta \\ &= \frac{1}{2}\left(-7\theta + 12\cos\theta + \sin(2\theta) \right)\Bigg|_{\frac{7\pi}{6}}^{\frac{11\pi}{6}} \\ &= \frac{11\sqrt{3}}{2} - \frac{7\pi}{3} = 2.196 \end{aligned}$$

Notice that for this area the "outer" and "inner" function were opposite!

Let's do one final modification of this example.

Example 4 Determine the area that is inside both $r = 3 + 2\sin\theta$ and $r = 2$.

Solution
Here is the sketch for this example.



We are not going to be able to do this problem in the same fashion that we did the previous two. There is no set of limits that will allow us to enclose this area as we increase from one to the other. Remember that as we increase θ the area we're after must be enclosed. However, the only two ranges for θ that we can work with enclose the area from the previous two examples and not this region.

In this case however, that is not a major problem. There are two ways to do get the area in this problem. We'll take a look at both of them.

Solution 1
In this case let's notice that the circle is divided up into two portions and we're after the upper portion. Also notice that we found the area of the lower portion in Example 3. Therefore, the area is,

$$\begin{aligned}\text{Area} &= \text{Area of Circle} - \text{Area from Example 3}\\ &= \pi(2)^2 - 2.196\\ &= 10.370\end{aligned}$$

Solution 2
In this case we do pretty much the same thing except this time we'll think of the area as the other portion of the limacon than the portion that we were dealing with in Example 2. We'll also need to actually compute the area of the limacon in this case.

So, the area using this approach is then,

$$\begin{aligned}\text{Area} &= \text{Area of Limacon} - \text{Area from Example 2}\\ &= \int_0^{2\pi} \frac{1}{2}\left(3+2\sin\theta\right)^2 \, d\theta - 24.187\\ &= \int_0^{2\pi} \frac{1}{2}\left(9+12\sin\theta+4\sin^2\theta\right) d\theta - 24.187\\ &= \int_0^{2\pi} \frac{1}{2}\left(11+12\sin\theta-2\cos\left(2\theta\right)\right) d\theta - 24.187\\ &= \frac{1}{2}\left(11\theta-12\cos\left(\theta\right)-\sin\left(2\theta\right)\right)\Big|_0^{2\pi} - 24.187\\ &= 11\pi - 24.187\\ &= 10.370\end{aligned}$$

A slightly longer approach, but sometimes we are forced to take this longer approach.

As this last example has shown we will not be able to get all areas in polar coordinates straight from an integral.

Arc Length with Polar Coordinates

We now need to move into the Calculus II applications of integrals and how we do them in terms of polar coordinates. In this section we'll look at the arc length of the curve given by,

$$r = f\left(\theta\right) \qquad\qquad \alpha \le \theta \le \beta$$

where we also assume that the curve is traced out exactly once. Just as we did with the tangent lines in polar coordinates we'll first write the curve in terms of a set of parametric equations,

$$\begin{aligned} x &= r\cos\theta & \qquad y &= r\sin\theta \\ &= f\left(\theta\right)\cos\theta & &= f\left(\theta\right)\sin\theta \end{aligned}$$

and we can now use the parametric formula for finding the arc length.

We'll need the following derivatives for these computations.

$$\begin{aligned} \frac{dx}{d\theta} &= f'\left(\theta\right)\cos\theta - f\left(\theta\right)\sin\theta & \qquad \frac{dy}{d\theta} &= f'\left(\theta\right)\sin\theta + f\left(\theta\right)\cos\theta \\ &= \frac{dr}{d\theta}\cos\theta - r\sin\theta & &= \frac{dr}{d\theta}\sin\theta + r\cos\theta \end{aligned}$$

We'll need the following for our *ds*.

$$\begin{aligned} \left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2 &= \left(\frac{dr}{d\theta}\cos\theta - r\sin\theta\right)^2 + \left(\frac{dr}{d\theta}\sin\theta + r\cos\theta\right)^2 \\ &= \left(\frac{dr}{d\theta}\right)^2\cos^2\theta - 2r\frac{dr}{d\theta}\cos\theta\sin\theta + r^2\sin^2\theta \\ &\qquad + \left(\frac{dr}{d\theta}\right)^2\sin^2\theta + 2r\frac{dr}{d\theta}\cos\theta\sin\theta + r^2\cos^2\theta \\ &= \left(\frac{dr}{d\theta}\right)^2\left(\cos^2\theta + \sin^2\theta\right) + r^2\left(\cos^2\theta + \sin^2\theta\right) \\ &= r^2 + \left(\frac{dr}{d\theta}\right)^2 \end{aligned}$$

The arc length formula for polar coordinates is then,

$$L = \int ds$$

where,

$$ds = \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta$$

Let's work a quick example of this.

Example 1 Determine the length of $r=\theta \quad 0\le\theta\le 1$.

Solution
Okay, let's just jump straight into the formula since this is a fairly simple function.

$$L=\int_0^1 \sqrt{\theta^2+1}\,d\theta$$

We'll need to use a trig substitution here.

$$\theta=\tan x \qquad\qquad d\theta=\sec^2 x\,dx$$

$$\theta=0 \qquad 0=\tan x \qquad x=0$$

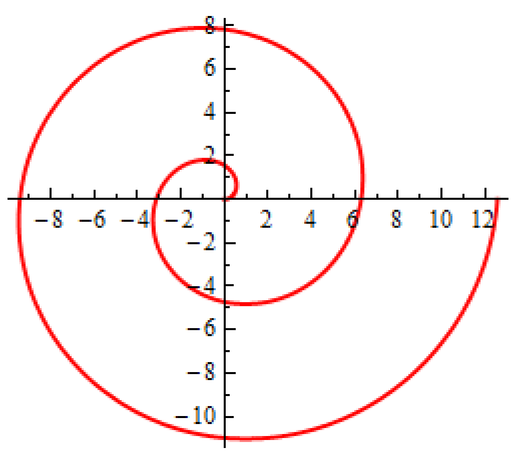
$$\theta=1 \qquad 1=\tan x \qquad x=\frac{\pi}{4}$$

$$\sqrt{\theta^2+1}=\sqrt{\tan^2 x+1}=\sqrt{\sec^2 x}=\left|\sec x\right|=\sec x$$

The arc length is then,

$$\begin{aligned} L&=\int_0^1 \sqrt{\theta^2+1}\,d\theta\\ &=\int_0^{\frac{\pi}{4}} \sec^3 x\,dx\\ &=\frac{1}{2}\left(\sec x\tan x+\ln\left|\sec x+\tan x\right|\right)\Big|_0^{\frac{\pi}{4}}\\ &=\frac{1}{2}\left(\sqrt{2}+\ln\left(1+\sqrt{2}\right)\right)\end{aligned}$$

Just as an aside before we leave this chapter. The polar equation $r=\theta$ is the equation of a spiral. Here is a quick sketch of $r=\theta$ for $0\le\theta\le 4\pi$.



Surface Area with Polar Coordinates

We will be looking at surface area in polar coordinates in this section. Note however that all we're going to do is give the formulas for the surface area since most of these integrals tend to be fairly difficult.

We want to find the surface area of the region found by rotating,

$$r = f\left(\theta\right) \qquad\qquad \alpha \le \theta \le \beta$$

about the x or y-axis.

As we did in the tangent and arc length sections we'll write the curve in terms of a set of parametric equations.

$$\begin{aligned} x &= r\cos\theta & \qquad\qquad y &= r\sin\theta \\ &= f\left(\theta\right)\cos\theta & &= f\left(\theta\right)\sin\theta \end{aligned}$$

If we now use the parametric formula for finding the surface area we'll get,

$$S = \int 2\pi y\, ds \qquad\qquad \text{rotation about } x-\text{axis}$$

$$S = \int 2\pi x\, ds \qquad\qquad \text{rotation about } y-\text{axis}$$

where,

$$ds = \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta \qquad\qquad r = f\left(\theta\right), \quad \alpha \le \theta \le \beta$$

Note that because we will pick up a $d\theta$ from the ds we'll need to substitute one of the parametric equations in for x or y depending on the axis of rotation. This will often mean that the integrals will be somewhat unpleasant.

Arc Length and Surface Area Revisited

We won't be working any examples in this section. This section is here solely for the purpose of summarizing up all the arc length and surface area problems.

Over the course of the last two chapters the topic of arc length and surface area has arisen many times and each time we got a new formula out of the mix. Students often get a little overwhelmed with all the formulas.

However, there really aren't as many formulas as it might seem at first glance. There is exactly one arc length formula and exactly two surface area formulas. These are,

$$L = \int ds$$

$$S = \int 2\pi y \, ds \qquad \text{rotation about } x-\text{axis}$$

$$S = \int 2\pi x \, ds \qquad \text{rotation about } y-\text{axis}$$

The problems arise because we have quite a few *ds*'s that we can use. Again students often have trouble deciding which one to use. The examples/problems usually suggest the correct one to use however. Here is a complete listing of all the *ds*'s that we've seen and when they are used.

$$ds = \sqrt{1+\left(\frac{dy}{dx}\right)^2}\, dx \qquad \text{if } y = f(x),\ a \le x \le b$$

$$ds = \sqrt{1+\left(\frac{dx}{dy}\right)^2}\, dy \qquad \text{if } x = h(y),\ c \le y \le d$$

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}\, dt \qquad \text{if } x = f(t), y = g(t),\ \alpha \le t \le \beta$$

$$ds = \sqrt{r^2+\left(\frac{dr}{d\theta}\right)^2}\, d\theta \qquad \text{if } r = f(\theta),\quad \alpha \le \theta \le \beta$$

Depending on the form of the function we can quickly tell which *ds* to use.

There is only one other thing to worry about in terms of the surface area formula. The *ds* will introduce a new differential to the integral. Before integrating make sure all the variables are in terms of this new differential. For example if we have parametric equations we'll use the third *ds* and then we'll need to make sure and substitute for the *x* or *y* depending on which axis we rotate about to get everything in terms of *t*.

Likewise, if we have a function in the form $x = h(y)$ then we'll use the second *ds* and if the rotation is about the *y*-axis we'll need to substitute for the *x* in the integral. On the other hand if we rotate about the *x*-axis we won't need to do a substitution for the *y*.

Keep these rules in mind and you'll always be able to determine which formula to use and how to correctly do the integral.

Sequences and Series

Introduction

In this chapter we'll be taking a look at sequences and (infinite) series. Actually, this chapter will deal almost exclusively with series. However, we also need to understand some of the basics of sequences in order to properly deal with series. We will therefore, spend a little time on sequences as well.

Series is one of those topics that many students don't find all that useful. To be honest, many students will never see series outside of their calculus class. However, series do play an important role in the field of ordinary differential equations and without series large portions of the field of partial differential equations would not be possible.

In other words, series is an important topic even if you won't ever see any of the applications. Most of the applications are beyond the scope of most Calculus courses and tend to occur in classes that many students don't take. So, as you go through this material keep in mind that these do have applications even if we won't really be covering many of them in this class.

Here is a list of topics in this chapter.

Sequences – We will start the chapter off with a brief discussion of sequences. This section will focus on the basic terminology and convergence of sequences

More on Sequences – Here we will take a quick look about monotonic and bounded sequences.

Series – The Basics – In this section we will discuss some of the basics of infinite series.

Series – Convergence/Divergence – Most of this chapter will be about the convergence/divergence of a series so we will give the basic ideas and definitions in this section.

Series – Special Series – We will look at the Geometric Series, Telescoping Series, and Harmonic Series in this section.

Integral Test – Using the Integral Test to determine if a series converges or diverges.

Comparison Test/Limit Comparison Test – Using the Comparison Test and Limit Comparison Tests to determine if a series converges or diverges.

Alternating Series Test – Using the Alternating Series Test to determine if a series converges or diverges.

Absolute Convergence – A brief discussion on absolute convergence and how it differs from convergence.

Sequences

Let's start off this section with a discussion of just what a sequence is. A sequence is nothing more than a list of numbers written in a specific order. The list may or may not have an infinite number of terms in them although we will be dealing exclusively with infinite sequences in this class. General sequence terms are denoted as follows,

$$\begin{aligned}
&a_1 - \text{first term}\\
&a_2 - \text{second term}\\
&\qquad\vdots\\
&a_n - n^{th} \text{ term}\\
&a_{n+1} - (n+1)^{\text{st}} \text{ term}\\
&\qquad\vdots
\end{aligned}$$

Because we will be dealing with infinite sequences each term in the sequence will be followed by another term as noted above. In the notation above we need to be very careful with the subscripts. The subscript of $n+1$ denotes the next term in the sequence and NOT one plus the n^{th} term! In other words,

$$a_{n+1} \neq a_n + 1$$

so be very careful when writing subscripts to make sure that the "+1" doesn't migrate out of the subscript! This is an easy mistake to make when you first start dealing with this kind of thing.

There is a variety of ways of denoting a sequence. Each of the following are equivalent ways of denoting a sequence.

$$\left\{a_1, a_2, \ldots, a_n, a_{n+1}, \ldots\right\} \qquad \left\{a_n\right\} \qquad \left\{a_n\right\}_{n=1}^{\infty}$$

In the second and third notations above a_n is usually given by a formula.

A couple of notes are now in order about these notations. First, note the difference between the second and third notations above. If the starting point is not important or is implied in some way by the problem it is often not written down as we did in the third notation. Next, we used a starting point of $n=1$ in the third notation only so we could write one down. There is absolutely no reason to believe that a sequence will start at $n=1$. A sequence will start where ever it needs to start.

Let's take a look at a couple of sequences.

Example 1 Write down the first few terms of each of the following sequences.

(a) $\left\{ \dfrac{n+1}{n^2} \right\}_{n=1}^{\infty}$ [Solution]

(b) $\left\{ \dfrac{(-1)^{n+1}}{2^n} \right\}_{n=0}^{\infty}$ [Solution]

(c) $\left\{ b_n \right\}_{n=1}^{\infty}$, where $b_n = n^{th}$ digit of π [Solution]

Solution

(a) $\left\{ \dfrac{n+1}{n^2} \right\}_{n=1}^{\infty}$

To get the first few sequence terms here all we need to do is plug in values of n into the formula given and we'll get the sequence terms.

$$\left\{ \frac{n+1}{n^2} \right\}_{n=1}^{\infty} = \left\{ \underbrace{2}_{n=1}, \underbrace{\frac{3}{4}}_{n=2}, \underbrace{\frac{4}{9}}_{n=3}, \underbrace{\frac{5}{16}}_{n=4}, \underbrace{\frac{6}{25}}_{n=5}, \ldots \right\}$$

Note the inclusion of the "…" at the end! This is an important piece of notation as it is the only thing that tells us that the sequence continues on and doesn't terminate at the last term.

[Return to Problems]

(b) $\left\{ \dfrac{(-1)^{n+1}}{2^n} \right\}_{n=0}^{\infty}$

This one is similar to the first one. The main difference is that this sequence doesn't start at $n = 1$.

$$\left\{ \frac{(-1)^{n+1}}{2^n} \right\}_{n=0}^{\infty} = \left\{ -1, \frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, -\frac{1}{16}, \ldots \right\}$$

Note that the terms in this sequence alternate in signs. Sequences of this kind are sometimes called alternating sequences.

[Return to Problems]

(c) $\left\{ b_n \right\}_{n=1}^{\infty}$, where $b_n = n^{th}$ digit of π

This sequence is different from the first two in the sense that it doesn't have a specific formula for each term. However, it does tell us what each term should be. Each term should be the n^{th} digit of π. So we know that $\pi = 3.14159265359\ldots$

The sequence is then,

$$\left\{ 3, 1, 4, 1, 5, 9, 2, 6, 5, 3, 5, \ldots \right\}$$

[Return to Problems]

In the first two parts of the previous example note that we were really treating the formulas as functions that can only have integers plugged into them. Or,

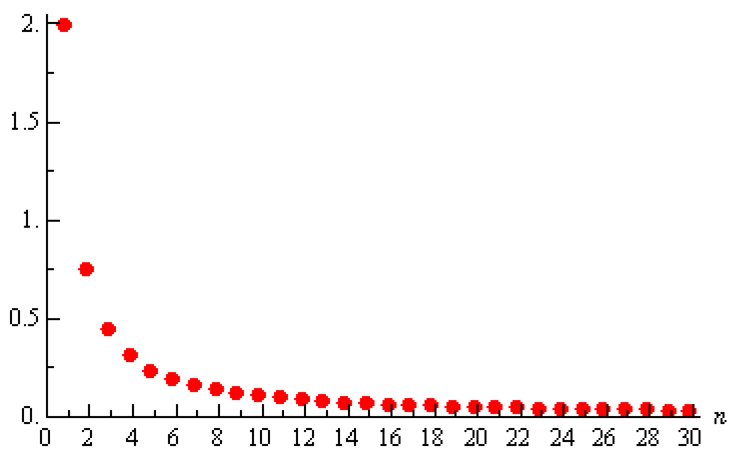
$$f(n) = \frac{n+1}{n^2} \qquad\qquad g(n) = \frac{(-1)^{n+1}}{2^n}$$

This is an important idea in the study of sequences (and series). Treating the sequence terms as function evaluations will allow us to do many things with sequences that couldn't do otherwise. Before delving further into this idea however we need to get a couple more ideas out of the way.

First we want to think about "graphing" a sequence. To graph the sequence $\{a_n\}$ we plot the points (n, a_n) as n ranges over all possible values on a graph. For instance, let's graph the sequence $\left\{\frac{n+1}{n^2}\right\}_{n=1}^{\infty}$. The first few points on the graph are,

$$(1,2),\ \left(2,\frac{3}{4}\right),\ \left(3,\frac{4}{9}\right),\ \left(4,\frac{5}{16}\right),\ \left(5,\frac{6}{25}\right),\ \ldots$$

The graph, for the first 30 terms of the sequence, is then,



This graph leads us to an important idea about sequences. Notice that as n increases the sequence terms in our sequence, in this case, get closer and closer to zero. We then say that zero is the **limit** (or sometimes the **limiting value**) of the sequence and write,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n+1}{n^2} = 0$$

This notation should look familiar to you. It is the same notation we used when we talked about the limit of a function. In fact, if you recall, we said earlier that we could think of sequences as functions in some way and so this notation shouldn't be too surprising.

Using the ideas that we developed for limits of functions we can write down the following *working definition* for limits of sequences.

***Working* Definition of Limit**

1. We say that
$$\lim_{n\to\infty} a_n = L$$
if we can make a_n as close to L as we want for all sufficiently large n. In other words, the value of the a_n's approach L as n approaches infinity.

2. We say that
$$\lim_{n\to\infty} a_n = \infty$$
if we can make a_n as large as we want for all sufficiently large n. Again, in other words, the value of the a_n's get larger and larger without bound as n approaches infinity.

3. We say that
$$\lim_{n\to\infty} a_n = -\infty$$
if we can make a_n as large and negative as we want for all sufficiently large n. Again, in other words, the value of the a_n's are negative and get larger and larger without bound as n approaches infinity.

The working definitions of the various sequence limits are nice in that they help us to visualize what the limit actually is. Just like with limits of functions however, there is also a precise definition for each of these limits. Let's give those before proceeding

***Precise* Definition of Limit**

1. We say that $\lim_{n\to\infty} a_n = L$ if for every number $\varepsilon > 0$ there is an integer N such that
$$\left|a_n - L\right| < \varepsilon \qquad \text{whenever} \qquad n > N$$

2. We say that $\lim_{n\to\infty} a_n = \infty$ if for every number $M > 0$ there is an integer N such that
$$a_n > M \qquad \text{whenever} \qquad n > N$$

3. We say that $\lim_{n\to\infty} a_n = -\infty$ if for every number $M < 0$ there is an integer N such that
$$a_n < M \qquad \text{whenever} \qquad n > N$$

We won't be using the precise definition often, but it will show up occasionally.

Note that both definitions tell us that in order for a limit to exist and have a finite value all the sequence terms must be getting closer and closer to that finite value as n increases.

Now that we have the definitions of the limit of sequences out of the way we have a bit of terminology that we need to look at. If $\lim_{n\to\infty} a_n$ exists and is finite we say that the sequence is **convergent**. If $\lim_{n\to\infty} a_n$ doesn't exist or is infinite we say the sequence **diverges**. Note that sometimes we will say the sequence **diverges to** ∞ if $\lim_{n\to\infty} a_n = \infty$ and if $\lim_{n\to\infty} a_n = -\infty$ we will sometimes say that the sequence **diverges to** $-\infty$.

Get used to the terms "convergent" and "divergent" as we'll be seeing them quite a bit throughout this chapter.

So just how do we find the limits of sequences? Most limits of most sequences can be found using one of the following theorems.

Theorem 1

Given the sequence $\left\{a_n\right\}$ if we have a function $f\left(x\right)$ such that $f\left(n\right)=a_n$ and $\lim\limits_{x\to\infty} f\left(x\right)=L$ then $\lim\limits_{n\to\infty} a_n = L$

This theorem is basically telling us that we take the limits of sequences much like we take the limit of functions. In fact, in most cases we'll not even really use this theorem by explicitly writing down a function. We will more often just treat the limit as if it were a limit of a function and take the limit as we always did back in Calculus I when we were taking the limits of functions.

So, now that we know that taking the limit of a sequence is nearly identical to taking the limit of a function we also know that all the properties from the limits of functions will also hold.

Properties

1. $\lim\limits_{n\to\infty}\left(a_n \pm b_n\right)=\lim\limits_{n\to\infty} a_n \pm \lim\limits_{n\to\infty} b_n$

2. $\lim\limits_{n\to\infty} c a_n = c\lim\limits_{n\to\infty} a_n$

3. $\lim\limits_{n\to\infty}\left(a_n\, b_n\right)=\left(\lim\limits_{n\to\infty} a_n\right)\left(\lim\limits_{n\to\infty} b_n\right)$

4. $\lim\limits_{n\to\infty}\dfrac{a_n}{b_n}=\dfrac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n}$, provided $\lim\limits_{n\to\infty} b_n \neq 0$

5. $\lim\limits_{n\to\infty} a_n^p = \left[\lim\limits_{n\to\infty} a_n\right]^p$ provided $a_n \geq 0$

These properties can be proved using Theorem 1 above and the function limit properties we saw in Calculus I or we can prove them directly using the precise definition of a limit using nearly identical proofs of the function limit properties.

Next, just as we had a Squeeze Theorem for function limits we also have one for sequences and it is pretty much identical to the function limit version.

Squeeze Theorem for Sequences

If $a_n \leq c_n \leq b_n$ for all $n > N$ for some N and $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n = L$ then $\lim\limits_{n\to\infty} c_n = L$.

Note that in this theorem the "for all $n > N$ for some N" is really just telling us that we need to have $a_n \le c_n \le b_n$ for all sufficiently large n, but if it isn't true for the first few n that won't invalidate the theorem.

As we'll see not all sequences can be written as functions that we can actually take the limit of. This will be especially true for sequences that alternate in signs. While we can always write these sequence terms as a function we simply don't know how to take the limit of a function like that. The following theorem will help with some of these sequences.

Theorem 2

If $\lim\limits_{n\to\infty} |a_n| = 0$ then $\lim\limits_{n\to\infty} a_n = 0$.

Note that in order for this theorem to hold the limit MUST be zero and it won't work for a sequence whose limit is not zero. This theorem is easy enough to prove so let's do that.

Proof of Theorem 2

The main thing to this proof is to note that,

$$-|a_n| \le a_n \le |a_n|$$

Then note that,

$$\lim_{n\to\infty}\left(-|a_n|\right) = -\lim_{n\to\infty}|a_n| = 0$$

We then have $\lim\limits_{n\to\infty}\left(-|a_n|\right) = \lim\limits_{n\to\infty}|a_n| = 0$ and so by the Squeeze Theorem we must also have,

$$\lim_{n\to\infty} a_n = 0$$

■

The next theorem is a useful theorem giving the convergence/divergence and value (for when it's convergent) of a sequence that arises on occasion.

Theorem 3

The sequence $\left\{r^n\right\}_{n=0}^{\infty}$ converges if $-1 < r \le 1$ and diverges for all other values of r. Also,

$$\lim_{n\to\infty} r^n = \begin{cases} 0 & \text{if } -1 < r < 1 \\ 1 & \text{if } r = 1 \end{cases}$$

Here is a quick (well not so quick, but definitely simple) partial proof of this theorem.

Partial Proof of Theorem 3

We'll do this by a series of cases although the last case will not be completely proven.

Case 1 : $r > 1$

We know from Calculus I that $\lim\limits_{x\to\infty} r^x = \infty$ if $r > 1$ and so by Theorem 1 above we also know that $\lim_{n\to\infty} r^n = \infty$ and so the sequence diverges if $r > 1$.

Case 2 : $r = 1$
In this case we have,

$$\lim_{n\to\infty} r^n = \lim_{n\to\infty} 1^n = \lim_{n\to\infty} 1 = 1$$

So, the sequence converges for $r = 1$ and in this case its limit is 1.

Case 3 : $0 < r < 1$
We know from Calculus I that $\lim_{x\to\infty} r^x = 0$ if $0 < r < 1$ and so by Theorem 1 above we also know that $\lim_{n\to\infty} r^n = 0$ and so the sequence converges if $0 < r < 1$ and in this case its limit is zero.

Case 4 : $r = 0$
In this case we have,

$$\lim_{n\to\infty} r^n = \lim_{n\to\infty} 0^n = \lim_{n\to\infty} 0 = 0$$

So, the sequence converges for $r = 0$ and in this case its limit is zero.

Case 5 : $-1 < r < 0$
First let's note that if $-1 < r < 0$ then $0 < |r| < 1$ then by Case 3 above we have,

$$\lim_{n\to\infty} |r^n| = \lim_{n\to\infty} |r|^n = 0$$

Theorem 2 above now tells us that we must also have, $\lim_{n\to\infty} r^n = 0$ and so if $-1 < r < 0$ the sequence converges and has a limit of 0.

Case 6 : $r = -1$
In this case the sequence is,

$$\left\{ r^n \right\}_{n=0}^{\infty} = \left\{ (-1)^n \right\}_{n=0}^{\infty} = \left\{ 1, -1, 1, -1, 1, -1, 1, -1, \ldots \right\}_{n=0}^{\infty}$$

and hopefully it is clear that $\lim_{n\to\infty} (-1)^n$ doesn't exist. Recall that in order of this limit to exist the terms must be approaching a single value as *n* increases. In this case however the terms just alternate between 1 and -1 and so the limit does not exist.

So, the sequence diverges for $r = -1$.

Case 7 : $r < -1$
In this case we're not going to go through a complete proof. Let's just see what happens if we let $r = -2$ for instance. If we do that the sequence becomes,

$$\left\{ r^n \right\}_{n=0}^{\infty} = \left\{ (-2)^n \right\}_{n=0}^{\infty} = \left\{ 1, -2, 4, -8, 16, -32, \ldots \right\}_{n=0}^{\infty}$$

So, if $r = -2$ we get a sequence of terms whose values alternate in sign and get larger and larger and so $\lim\limits_{n\to\infty}(-2)^n$ doesn't exist. It does not settle down to a single value as *n* increases nor do the terms ALL approach infinity. So, the sequence diverges for $r=-2$.

We could do something similar for any value of *r* such that $r<-1$ and so the sequence diverges for $r<-1$.
■

Let's take a look at a couple of examples of limits of sequences.

Example 2 Determine if the following sequences converge or diverge. If the sequence converges determine its limit.

(a) $\left\{\dfrac{3n^2-1}{10n+5n^2}\right\}_{n=2}^{\infty}$ [Solution]

(b) $\left\{\dfrac{\mathbf{e}^{2n}}{n}\right\}_{n=1}^{\infty}$ [Solution]

(c) $\left\{\dfrac{(-1)^n}{n}\right\}_{n=1}^{\infty}$ [Solution]

(d) $\left\{(-1)^n\right\}_{n=0}^{\infty}$ [Solution]

Solution

(a) $\left\{\dfrac{3n^2-1}{10n+5n^2}\right\}_{n=2}^{\infty}$

In this case all we need to do is recall the method that was developed in Calculus I to deal with the limits of rational functions. See the Limits At Infinity, Part I section of my Calculus I notes for a review of this if you need to.

To do a limit in this form all we need to do is factor from the numerator and denominator the largest power of *n*, cancel and then take the limit.

$$\lim_{n\to\infty}\frac{3n^2-1}{10n+5n^2}=\lim_{n\to\infty}\frac{n^2\left(3-\frac{1}{n^2}\right)}{n^2\left(\frac{10}{n}+5\right)}=\lim_{n\to\infty}\frac{3-\frac{1}{n^2}}{\frac{10}{n}+5}=\frac{3}{5}$$

So the sequence converges and its limit is $\frac{3}{5}$.

[Return to Problems]

(b) $\left\{ \dfrac{\mathbf{e}^{2n}}{n} \right\}_{n=1}^{\infty}$

We will need to be careful with this one. We will need to use L'Hospital's Rule on this sequence. The problem is that L'Hospital's Rule only works on functions and not on sequences. Normally this would be a problem, but we've got Theorem 1 from above to help us out. Let's define

$$f(x) = \frac{\mathbf{e}^{2x}}{x}$$

and note that,

$$f(n) = \frac{\mathbf{e}^{2n}}{n}$$

Theorem 1 says that all we need to do is take the limit of the function.

$$\lim_{n \to \infty} \frac{\mathbf{e}^{2n}}{n} = \lim_{x \to \infty} \frac{\mathbf{e}^{2x}}{x} = \lim_{x \to \infty} \frac{2\mathbf{e}^{2x}}{1} = \infty$$

So, the sequence in this part diverges (to ∞).

More often than not we just do L'Hospital's Rule on the sequence terms without first converting to *x*'s since the work will be identical regardless of whether we use *x* or *n*. However, we really should remember that technically we can't do the derivatives while dealing with sequence terms.

[Return to Problems]

(c) $\left\{ \dfrac{(-1)^n}{n} \right\}_{n=1}^{\infty}$

We will also need to be careful with this sequence. We might be tempted to just say that the limit of the sequence terms is zero (and we'd be correct). However, technically we can't take the limit of sequences whose terms alternate in sign, because we don't know how to do limits of functions that exhibit that same behavior. Also, we want to be very careful to not rely too much on intuition with these problems. As we will see in the next section, and in later sections, our intuition can lead us astray in these problem if we aren't careful.

So, let's work this one by the book. We will need to use Theorem 2 on this problem. To this we'll first need to compute,

$$\lim_{n \to \infty} \left| \frac{(-1)^n}{n} \right| = \lim_{n \to \infty} \frac{1}{n} = 0$$

Therefore, since the limit of the sequence terms with absolute value bars on them goes to zero we know by Theorem 2 that,

$$\lim_{n \to \infty} \frac{(-1)^n}{n} = 0$$

which also means that the sequence converges to a value of zero.

[Return to Problems]

(d) $\left\{(-1)^n\right\}_{n=0}^{\infty}$

For this theorem note that all we need to do is realize that this is the sequence in Theorem 3 above using $r=-1$. So, by Theorem 3 this sequence diverges.

[Return to Problems]

We now need to give a warning about misusing Theorem 2. Theorem 2 only works if the limit is zero. If the limit of the absolute value of the sequence terms is not zero then the theorem will not hold. The last part of the previous example is a good example of this (and in fact this warning is the whole reason that part is there). Notice that

$$\lim_{n\to\infty}\left|(-1)^n\right| = \lim_{n\to\infty} 1 = 1$$

and yet, $\lim_{n\to\infty}(-1)^n$ doesn't even exist let alone equal 1. So, be careful using this Theorem 2. You must always remember that it only works if the limit is zero.

Before moving onto the next section we need to give one more theorem that we'll need for a proof down the road.

Theorem 4

For the sequence $\{a_n\}$ if both $\lim_{n\to\infty} a_{2n} = L$ and $\lim_{n\to\infty} a_{2n+1} = L$ then $\{a_n\}$ is convergent and $\lim_{n\to\infty} a_n = L$.

Proof of Theorem 4

Let $\varepsilon > 0$.

Then since $\lim_{n\to\infty} a_{2n} = L$ there is an $N_1 > 0$ such that if $n > N_1$ we know that,

$$\left|a_{2n} - L\right| < \varepsilon$$

Likewise, because $\lim_{n\to\infty} a_{2n+1} = L$ there is an $N_2 > 0$ such that if $n > N_2$ we know that,

$$\left|a_{2n+1} - L\right| < \varepsilon$$

Now, let $N = \max\{2N_1, 2N_2+1\}$ and let $n > N$. Then either $a_n = a_{2k}$ for some $k > N_1$ or $a_n = a_{2k+1}$ for some $k > N_2$ and so in either case we have that,

$$\left|a_n - L\right| < \varepsilon$$

Therefore, $\lim_{n\to\infty} a_n = L$ and so $\{a_n\}$ is convergent.

■

More on Sequences

In the previous section we introduced the concept of a sequence and talked about limits of sequences and the idea of convergence and divergence for a sequence. In this section we want to take a quick look at some ideas involving sequences.

Let's start off with some terminology and definitions.

Given any sequence $\{a_n\}$ we have the following.

1. We call the sequence **increasing** if $a_n < a_{n+1}$ for every *n*.
2. We call the sequence **decreasing** if $a_n > a_{n+1}$ for every *n*.
3. If $\{a_n\}$ is an increasing sequence or $\{a_n\}$ is a decreasing sequence we call it **monotonic**.
4. If there exists a number *m* such that $m \le a_n$ for every *n* we say the sequence is **bounded below**. The number *m* is sometimes called a **lower bound** for the sequence.
5. If there exists a number *M* such that $a_n \le M$ for every *n* we say the sequence is **bounded above**. The number *M* is sometimes called an **upper bound** for the sequence.
6. If the sequence is both bounded below and bounded above we call the sequence **bounded**.

Note that in order for a sequence to be increasing or decreasing it must be increasing/decreasing for every *n*. In other words, a sequence that increases for three terms and then decreases for the rest of the terms is NOT a decreasing sequence! Also note that a monotonic sequence must always increase or it must always decrease.

Before moving on we should make a quick point about the bounds for a sequence that is bounded above and/or below. We'll make the point about lower bounds, but we could just as easily make it about upper bounds.

A sequence is bounded below if we can find any number *m* such that $m \le a_n$ for every *n*. Note however that if we find one number *m* to use for a lower bound then any number smaller than *m* will also be a lower bound. Also, just because we find one lower bound that doesn't mean there won't be a "better" lower bound for the sequence than the one we found. In other words, there are an infinite number of lower bounds for a sequence that is bounded below, some will be better than others. In my class all that I'm after will be a lower bound. I don't necessarily need the best lower bound, just a number that will be a lower bound for the sequence.

Let's take a look at a couple of examples.

Example 1 Determine if the following sequences are monotonic and/or bounded.

(a) $\left\{-n^2\right\}_{n=0}^{\infty}$ [Solution]

(b) $\left\{\left(-1\right)^{n+1}\right\}_{n=1}^{\infty}$ [Solution]

(c) $\left\{\dfrac{2}{n^2}\right\}_{n=5}^{\infty}$ [Solution]

Solution

(a) $\left\{-n^2\right\}_{n=0}^{\infty}$

This sequence is a decreasing sequence (and hence monotonic) because,

$$-n^2 > -\left(n+1\right)^2$$

for every *n*.

Also, since the sequence terms will be either zero or negative this sequence is bounded above. We can use any positive number or zero as the bound, *M*, however, it's standard to choose the smallest possible bound if we can and it's a nice number. So, we'll choose $M = 0$ since,

$$-n^2 \le 0 \qquad \text{for every } n$$

This sequence is not bounded below however since we can always get below any potential bound by taking *n* large enough. Therefore, while the sequence is bounded above it is not bounded.

As a side note we can also note that this sequence diverges (to $-\infty$ if we want to be specific).

[Return to Problems]

(b) $\left\{\left(-1\right)^{n+1}\right\}_{n=1}^{\infty}$

The sequence terms in this sequence alternate between 1 and -1 and so the sequence is neither an increasing sequence or a decreasing sequence. Since the sequence is neither an increasing nor decreasing sequence it is not a monotonic sequence.

The sequence is bounded however since it is bounded above by 1 and bounded below by -1.

Again, we can note that this sequence is also divergent.

[Return to Problems]

(c) $\left\{\dfrac{2}{n^2}\right\}_{n=5}^{\infty}$

This sequence is a decreasing sequence (and hence monotonic) since,

$$\frac{2}{n^2} > \frac{2}{\left(n+1\right)^2}$$

The terms in this sequence are all positive and so it is bounded below by zero. Also, since the sequence is a decreasing sequence the first sequence term will be the largest and so we can see that the sequence will also be bounded above by $\frac{2}{25}$. Therefore, this sequence is bounded.

We can also take a quick limit and note that this sequence converges and its limit is zero.

[Return to Problems]

Now, let's work a couple more examples that are designed to make sure that we don't get too used to relying on our intuition with these problems. As we noted in the previous section our intuition can often lead us astray with some of the concepts we'll be looking at in this chapter.

Example 2 Determine if the following sequences are monotonic and/or bounded.

(a) $\left\{\dfrac{n}{n+1}\right\}_{n=1}^{\infty}$ [Solution]

(b) $\left\{\dfrac{n^3}{n^4+10000}\right\}_{n=0}^{\infty}$ [Solution]

Solution

(a) $\left\{\dfrac{n}{n+1}\right\}_{n=1}^{\infty}$

We'll start with the bounded part of this example first and then come back and deal with the increasing/decreasing question since that is where students often make mistakes with this type of sequence.

First, *n* is positive and so the sequence terms are all positive. The sequence is therefore bounded below by zero. Likewise each sequence term is the quotient of a number divided by a larger number and so is guaranteed to be less than one. The sequence is then bounded above by one. So, this sequence is bounded.

Now let's think about the monotonic question. First, students will often make the mistake of assuming that because the denominator is larger the quotient must be decreasing. This will not always be the case and in this case we would be wrong. This sequence is increasing as we'll see.

To determine the increasing/decreasing nature of this sequence we will need to resort to Calculus I techniques. First consider the following function and its derivative.

$$f(x)=\frac{x}{x+1} \qquad\qquad f'(x)=\frac{1}{(x+1)^2}$$

We can see that the first derivative is always positive and so from Calculus I we know that the function must then be an increasing function. So, how does this help us? Notice that,

$$f(n)=\frac{n}{n+1}=a_n$$

Therefore because $n<n+1$ and $f(x)$ is increasing we can also say that,

$$a_n = \frac{n}{n+1} = f(n) < f(n+1) = \frac{n+1}{n+1+1} = a_{n+1} \qquad \Rightarrow \qquad a_n < a_{n+1}$$

In other words, the sequence must be increasing.

Note that now that we know the sequence is an increasing sequence we can get a better lower bound for the sequence. Since the sequence is increasing the first term in the sequence must be the smallest term and so since we are starting at $n = 1$ we could also use a lower bound of $\frac{1}{2}$ for this sequence. It is important to remember that any number that is always less than or equal to all the sequence terms can be a lower bound. Some are better than others however.

A quick limit will also tell us that this sequence converges with a limit of 1.

Before moving on to the next part there is a natural question that many students will have at this point. Why did we use Calculus to determine the increasing/decreasing nature of the sequence when we could have just plugged in a couple of *n*'s and quickly determined the same thing?

The answer to this question is the next part of this example!

[Return to Problems]

(b) $\left\{ \dfrac{n^3}{n^4 + 10000} \right\}_{n=0}^{\infty}$

This is a messy looking sequence, but it needs to be in order to make the point of this part.

First, notice that, as with the previous part, the sequence terms are all positive and will all be less than one (since the numerator is guaranteed to be less than the denominator) and so the sequence is bounded.

Now, let's move on to the increasing/decreasing question. As with the last problem, many students will look at the exponents in the numerator and denominator and determine based on that that sequence terms must decrease.

This however, isn't a decreasing sequence. Let's take a look at the first few terms to see this.

$$a_1 = \frac{1}{10001} \approx 0.00009999 \qquad a_2 = \frac{1}{1252} \approx 0.0007987$$

$$a_3 = \frac{27}{10081} \approx 0.005678 \qquad a_4 = \frac{4}{641} \approx 0.006240$$

$$a_5 = \frac{1}{85} \approx 0.011756 \qquad a_6 = \frac{27}{1412} \approx 0.019122$$

$$a_7 = \frac{343}{12401} \approx 0.02766 \qquad a_8 = \frac{32}{881} \approx 0.03632$$

$$a_9 = \frac{729}{16561} \approx 0.04402 \qquad a_{10} = \frac{1}{20} = 0.05$$

The first 10 terms of this sequence are all increasing and so clearly the sequence can't be a decreasing sequence. Recall that a sequence can only be decreasing if ALL the terms are decreasing.

Now, we can't make another common mistake and assume that because the first few terms increase then whole sequence must also increase. If we did that we would also be mistaken as this is also not an increasing sequence.

This sequence is neither decreasing or increasing. The only sure way to see this is to do the Calculus I approach to increasing/decreasing functions.

In this case we'll need the following function and its derivative.

$$f(x) = \frac{x^3}{x^4 + 10000} \qquad\qquad f'(x) = \frac{-x^2\left(x^4 - 30000\right)}{\left(x^4 + 10000\right)^2}$$

This function will have the following three critical points,

$$x = 0, \quad x = \sqrt[4]{30000} \approx 13.1607, \quad x = -\sqrt[4]{30000} \approx -13.1607$$

Why critical points? Remember these are the only places where the function *may* change sign! Our sequence starts at $n = 0$ and so we can ignore the third one since it lies outside the values of n that we're considering. By plugging in some test values of x we can quickly determine that the derivative is positive for $0 < x < \sqrt[4]{30000} \approx 13.16$ and so the function is increasing in this range. Likewise, we can see that the derivative is negative for $x > \sqrt[4]{30000} \approx 13.16$ and so the function will be decreasing in this range.

So, our sequence will be increasing for $0 \le n \le 13$ and decreasing for $n \ge 13$. Therefore the function is not monotonic.

Finally, note that this sequence will also converge and has a limit of zero.

[Return to Problems]

So, as the last example has shown we need to be careful in making assumptions about sequences. Our intuition will often not be sufficient to get the correct answer and we can NEVER make assumptions about a sequence based on the value of the first few terms. As the last part has shown there are sequences which will increase or decrease for a few terms and then change direction after that.

Note as well that we said "first few terms" here, but it is completely possible for a sequence to decrease for the first 10,000 terms and then start increasing for the remaining terms. In other words, there is no "magical" value of n for which all we have to do is check up to that point and then we'll know what the whole sequence will do.

The only time that we'll be able to avoid using Calculus I techniques to determine the increasing/decreasing nature of a sequence is in sequences like part (c) of Example 1. In this case increasing n only changed (in fact increased) the denominator and so we were able to determine the behavior of the sequence based on that.

In Example 2 however, increasing *n* increased both the denominator and the numerator. In cases like this there is no way to determine which increase will "win out" and cause the sequence terms to increase or decrease and so we need to resort to Calculus I techniques to answer the question.

We'll close out this section with a nice theorem that we'll use in some of the proofs later in this chapter.

Theorem

If $\left\{a_n\right\}$ is bounded and monotonic then $\left\{a_n\right\}$ is convergent.

Be careful to not misuse this theorem. It does not say that if a sequence is not bounded and/or not monotonic that it is divergent. Example 2b is a good case in point. The sequence in that example was not monotonic but it does converge.

Note as well that we can make several variants of this theorem. If $\left\{a_n\right\}$ is bounded above and increasing then it converges and likewise if $\left\{a_n\right\}$ is bounded below and decreasing then it converges.

Series – The Basics

In this section we will introduce the topic that we will be discussing for the rest of this chapter. That topic is infinite series. So just what is an infinite series? Well, let's start with a sequence $\left\{a_n\right\}_{n=1}^{\infty}$ (note the $n=1$ is for convenience, it can be anything) and define the following,

$$\begin{aligned}
s_1 &= a_1 \\
s_2 &= a_1 + a_2 \\
s_3 &= a_1 + a_2 + a_3 \\
s_4 &= a_1 + a_2 + a_3 + a_4 \\
&\vdots \\
s_n &= a_1 + a_2 + a_3 + a_4 + \cdots + a_n = \sum_{i=1}^{n} a_i
\end{aligned}$$

The s_n are called **partial sums** and notice that they will form a sequence, $\left\{s_n\right\}_{n=1}^{\infty}$. Also recall that the Σ is used to represent this summation and called a variety of names. The most common names are : **series notation**, **summation notation**, and **sigma notation**.

You should have seen this notation, at least briefly, back when you saw the definition of a definite integral in Calculus I. If you need a quick refresher on summation notation see the review of summation notation in my Calculus I notes.

Now back to series. We want to take a look at the limit of the sequence of partial sums, $\left\{s_n\right\}_{n=1}^{\infty}$. Notationally we'll define,

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{i=1}^{n} a_i = \sum_{i=1}^{\infty} a_i$$

We will call $\displaystyle\sum_{i=1}^{\infty} a_i$ an **infinite series** and note that the series "starts" at $i=1$ because that is where our original sequence, $\left\{a_n\right\}_{n=1}^{\infty}$, started. Had our original sequence started at 2 then our infinite series would also have started at 2. The infinite series will start at the same value that the sequence of terms (as opposed to the sequence of partial sums) starts.

If the sequence of partial sums, $\left\{s_n\right\}_{n=1}^{\infty}$, is convergent and its limit is finite then we also call the infinite series, $\displaystyle\sum_{i=1}^{\infty} a_i$ **convergent** and if the sequence of partial sums is divergent then the infinite series is also called **divergent**.

Note that sometimes it is convenient to write the infinite series as,

$$\sum_{i=1}^{\infty} a_i = a_1 + a_2 + a_3 + \cdots + a_n + \cdots$$

We do have to be careful with this however. This implies that an infinite series is just an infinite sum of terms and as we'll see in the next section this is not really true.

In the next section we're going to be discussing in greater detail the value of an infinite series, provided it has one of course as well as the ideas of convergence and divergence.

This section is going to be devoted mostly to notational issues as well as making sure we can do some basic manipulations with infinite series so we are ready for them when we need to be able to deal with them in later sections.

First, we should note that in most of this chapter we will refer to infinite series as simply series. If we ever need to work with both infinite and finite series we'll be more careful with terminology, but in most sections we'll be dealing exclusively with infinite series and so we'll just call them series.

Now, in $\sum_{i=1}^{\infty} a_i$ the *i* is called the **index of summation** or just **index** for short and note that the letter we use to represent the index does not matter. So for example the following series are all the same. The only difference is the letter we've used for the index.

$$\sum_{i=0}^{\infty} \frac{3}{i^2+1} = \sum_{k=0}^{\infty} \frac{3}{k^2+1} = \sum_{n=0}^{\infty} \frac{3}{n^2+1} \quad etc.$$

It is important to again note that the index will start at whatever value the sequence of series terms starts at and this can literally be anything. So far we've used $n=0$ and $n=1$ but the index could have started anywhere. In fact, we will usually use $\sum a_n$ to represent an infinite series in which the starting point for the index is not important. When we drop the initial value of the index we'll also drop the infinity from the top so don't forget that it is still technically there.

We will be dropping the initial value of the index in quite a few facts and theorems that we'll be seeing throughout this chapter. In these facts/theorems the starting point of the series will not affect the result and so to simplify the notation and to avoid giving the impression that the starting point is important we will drop the index from the notation. Do not forget however, that there is a starting point and that this will be an infinite series.

Note however, that if we do put an initial value of the index on a series in a fact/theorem it is there because it really does need to be there.

Now that some of the notational issues are out of the way we need to start thinking about various ways that we can manipulate series.

We'll start this off with basic arithmetic with infinite series as we'll need to be able to do that on occasion. We have the following properties.

Properties

If $\sum a_n$ and $\sum b_n$ are both convergent series then,

6. $\sum ca_n$, where c is any number, is also convergent and

$$\sum ca_n = c\sum a_n$$

7. $\sum_{n=k}^{\infty} a_n \pm \sum_{n=k}^{\infty} b_n$ is also convergent and,

$$\sum_{n=k}^{\infty} a_n \pm \sum_{n=k}^{\infty} b_n = \sum_{n=k}^{\infty} \left(a_n \pm b_n\right).$$

The first property is simply telling us that we can always factor a multiplicative constant out of an infinite series and again recall that if we don't put in an initial value of the index that the series can start at any value. Also recall that in these cases we won't put an infinity at the top either.

The second property says that if we add/subtract series all we really need to do is add/subtract the series terms. Note as well that in order to add/subtract series we need to make sure that both have the same initial value of the index and the new series will also start at this value.

Before we move on to a different topic let's discuss multiplication of series briefly. We'll start both series at $n = 0$ for a later formula and then note that,

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) \neq \sum_{n=0}^{\infty}\left(a_n b_n\right)$$

To convince yourself that this isn't true consider the following product of two finite sums.

$$\left(2+x\right)\left(3-5x+x^2\right) = 6-7x-3x^2+x^3$$

Yeah, it was just the multiplication of two polynomials. Each is a finite sum and so it makes the point. In doing the multiplication we didn't just multiply the constant terms, then the x terms, *etc*. Instead we had to distribute the 2 through the second polynomial, then distribute the x through the second polynomial and finally combine like terms.

Multiplying infinite series (even though we said we can't think of an infinite series as an infinite sum) needs to be done in the same manner. With multiplication we're really asking us to do the following,

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \left(a_0 + a_1 + a_2 + a_3 + \cdots\right)\left(b_0 + b_1 + b_2 + b_3 + \cdots\right)$$

To do this multiplication we would have to distribute the a_0 through the second term, distribute the a_1 through, *etc* then combine like terms. This is pretty much impossible since both series have an infinite set of terms in them, however the following formula can be used to determine the product of two series.

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty} c_n \quad \text{where} \quad c_n = \sum_{i=0}^{n} a_i b_{n-i}$$

We also can't say a lot about the convergence of the product. Even if both of the original series are convergent it is possible for the product to be divergent. The reality is that multiplication of series is a somewhat difficult process and in general is avoided if possible. We will take a brief look at it towards the end of the chapter when we've got more work under our belt and we run across a situation where it might actually be what we want to do. Until then, don't worry about multiplying series.

The next topic that we need to discuss in this section is that of **index shift**. To be honest this is not a topic that we'll see all that often in this course. In fact, we'll use it once in the next section and then not use it again in all likelihood. Despite the fact that we won't use it much in this course doesn't mean however that it isn't used often in other classes where you might run across series. So, we will cover it briefly here so that you can say you've seen it.

The basic idea behind index shifts is to start a series at a different value for whatever the reason (and yes, there are legitimate reasons for doing that).

Consider the following series,

$$\sum_{n=2}^{\infty} \frac{n+5}{2^n}$$

Suppose that for some reason we wanted to start this series at $n = 0$, but we didn't want to change the value of the series. This means that we can't just change the $n = 2$ to $n = 0$ as this would add in two new terms to the series and thus change its value.

Performing an index shift is a fairly simple process to do. We'll start by defining a new index, say *i*, as follows,

$$i = n - 2$$

Now, when $n = 2$, we will get $i = 0$. Notice as well that if $n = \infty$ then $i = \infty - 2 = \infty$, so only the lower limit will change here. Next, we can solve this for *n* to get,

$$n = i + 2$$

We can now completely rewrite the series in terms of the index *i* instead of the index *n* simply by plugging in our equation for *n* in terms of *i*.

$$\sum_{n=2}^{\infty} \frac{n+5}{2^n} = \sum_{i=0}^{\infty} \frac{(i+2)+5}{2^{i+2}} = \sum_{i=0}^{\infty} \frac{i+7}{2^{i+2}}$$

To finish the problem out we'll recall that the letter we used for the index doesn't matter and so we'll change the final *i* back into an *n* to get,

$$\sum_{n=2}^{\infty} \frac{n+5}{2^n} = \sum_{n=0}^{\infty} \frac{n+7}{2^{n+2}}$$

To convince yourselves that these really are the same summation let's write out the first couple of terms for each of them,

$$\sum_{n=2}^{\infty} \frac{n+5}{2^n} = \frac{7}{2^2} + \frac{8}{2^3} + \frac{9}{2^4} + \frac{10}{2^5} + \cdots$$
$$\sum_{n=0}^{\infty} \frac{n+7}{2^{n+2}} = \frac{7}{2^2} + \frac{8}{2^3} + \frac{9}{2^4} + \frac{10}{2^5} + \cdots$$

So, sure enough the two series do have exactly the same terms.

There is actually an easier way to do an index shift. The method given above is the technically correct way of doing an index shift. However, notice in the above example we decreased the initial value of the index by 2 and all the *n*'s in the series terms increased by 2 as well.

This will always work in this manner. If we decrease the initial value of the index by a set amount then all the other *n*'s in the series term will increase by the same amount. Likewise, if we increase the initial value of the index by a set amount, then all the *n*'s in the series term will decrease by the same amount.

Let's do a couple of examples using this shorthand method for doing index shifts.

Example 1 Perform the following index shifts.

(a) Write $\sum_{n=1}^{\infty} ar^{n-1}$ as a series that starts at $n = 0$.

(b) Write $\sum_{n=1}^{\infty} \frac{n^2}{1-3^{n+1}}$ as a series that starts at $n = 3$.

Solution

(a) In this case we need to decrease the initial value by 1 and so the *n*'s (okay the single *n*) in the term must increase by 1 as well.

$$\sum_{n=1}^{\infty} ar^{n-1} = \sum_{n=0}^{\infty} ar^{(n+1)-1} = \sum_{n=0}^{\infty} ar^{n}$$

(b) For this problem we want to increase the initial value by 2 and so all the *n*'s in the series term must decrease by 2.

$$\sum_{n=1}^{\infty} \frac{n^2}{1-3^{n+1}} = \sum_{n=3}^{\infty} \frac{(n-2)^2}{1-3^{(n-2)+1}} = \sum_{n=3}^{\infty} \frac{(n-2)^2}{1-3^{n-1}}$$

The final topic in this section is again a topic that we'll not be seeing all that often in this class, although we will be seeing it more often than the index shifts. This final topic is really more about alternate ways to write series when the situation requires it.

Let's start with the following series and note that the $n = 1$ starting point is only for convenience since we need to start the series somewhere.

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + a_4 + a_5 + \cdots$$

Notice that if we ignore the first term the remaining terms will also be a series that will start at $n = 2$ instead of $n = 1$ So, we can rewrite the original series as follows,

$$\sum_{n=1}^{\infty} a_n = a_1 + \sum_{n=2}^{\infty} a_n$$

In this example we say that we've stripped out the first term.

We could have stripped out more terms if we wanted to. In the following series we've stripped out the first two terms and the first four terms respectively.

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \sum_{n=3}^{\infty} a_n$$
$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + a_4 + \sum_{n=5}^{\infty} a_n$$

Being able to strip out terms will, on occasion, simplify our work or allow us to reuse a prior result so it's an important idea to remember.

Notice that in the second example above we could have also denoted the four terms that we stripped out as a finite series as follows,

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + a_4 + \sum_{n=5}^{\infty} a_n = \sum_{n=1}^{4} a_n + \sum_{n=5}^{\infty} a_n$$

This is a convenient notation when we are stripping out a large number of terms or if we need to strip out an undetermined number of terms. In general, we can write a series as follows,

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{N} a_n + \sum_{n=N+1}^{\infty} a_n$$

We'll leave this section with an important warning about terminology. Don't get sequences and series confused! A sequence is a list of numbers written in a specific order while an infinite series is a limit of a sequence of finite series and hence, if it exists will be a single value.

So, once again, a sequence is a list of numbers while a series is a single number, provided it makes sense to even compute the series. Students will often confuse the two and try to use facts pertaining to one on the other. However, since they are different beasts this just won't work. There will be problems where we are using both sequences and series so we'll always have to remember that they are different.

Series – Convergence/Divergence

In the previous section we spent some time getting familiar with series and we briefly defined convergence and divergence. Before worrying about convergence and divergence of a series we wanted to make sure that we've started to get comfortable with the notation involved in series and some of the various manipulations of series that we will, on occasion, need to be able to do.

As noted in the previous section most of what we were doing there won't be done much in this chapter. So, it is now time to start talking about the convergence and divergence of a series as this will be a topic that we'll be dealing with to one extent or another in almost all of the remaining sections of this chapter.

So, let's recap just what an infinite series is and what it means for a series to be convergent or divergent. We'll start with a sequence $\left\{a_n\right\}_{n=1}^{\infty}$ and again note that we're starting the sequence at $n=1$ only for the sake of convenience and it can, in fact, be anything.

Next we define the partial sums of the series as,

$$\begin{aligned}
s_1 &= a_1 \\
s_2 &= a_1 + a_2 \\
s_3 &= a_1 + a_2 + a_3 \\
s_4 &= a_1 + a_2 + a_3 + a_4 \\
&\vdots \\
s_n &= a_1 + a_2 + a_3 + a_4 + \cdots + a_n = \sum_{i=1}^{n} a_i
\end{aligned}$$

and these form a new sequence, $\left\{s_n\right\}_{n=1}^{\infty}$.

An infinite series, or just series here since almost every series that we'll be looking at will be an infinite series, is then the limit of the partial sums. Or,

$$\sum_{i=1}^{\infty} a_i = \lim_{n\to\infty} s_n$$

If the sequence of partial sums is a convergent sequence (*i.e.* its limit exists and is finite) then the series is also called **convergent** and in this case if $\lim_{n\to\infty} s_n = s$ then, $\sum_{i=1}^{\infty} a_i = s$. Likewise, if the sequence of partial sums is a divergent sequence (*i.e.* its limit doesn't exist or is plus or minus infinity) then the series is also called **divergent**.

Let's take a look at some series and see if we can determine if they are convergent or divergent and see if we can determine the value of any convergent series we find.

Example 1 Determine if the following series is convergent or divergent. If it converges determine its value.

$$\sum_{n=1}^{\infty} n$$

Solution
To determine if the series is convergent we first need to get our hands on a formula for the general term in the sequence of partial sums.

$$s_n = \sum_{i=1}^{n} i$$

This is a known series and its value can be shown to be,

$$s_n = \sum_{i=1}^{n} i = \frac{n\left(n+1\right)}{2}$$

Don't worry if you didn't know this formula (I'd be surprised if anyone knew it…) as you won't be required to know it in my course.

So, to determine if the series is convergent we will first need to see if the sequence of partial sums,

$$\left\{ \frac{n\left(n+1\right)}{2} \right\}_{n=1}^{\infty}$$

is convergent or divergent. That's not terribly difficult in this case. The limit of the sequence terms is,

$$\lim_{n\to\infty} \frac{n\left(n+1\right)}{2} = \infty$$

Therefore, the sequence of partial sums diverges to ∞ and so the series also diverges.

So, as we saw in this example we had to know a fairly obscure formula in order to determine the convergence of this series. In general finding a formula for the general term in the sequence of partial sums is a very difficult process. In fact after the next section we'll not be doing much with the partial sums of series due to the extreme difficulty faced in finding the general formula. This also means that we'll not be doing much work with the value of series since in order to get the value we'll also need to know the general formula for the partial sums.

We will continue with a few more examples however, since this is technically how we determine convergence and the value of a series. Also, the remaining examples we'll be looking at in this section will lead us to a very important fact about the convergence of series.

So, let's take a look at a couple more examples.

Example 2 Determine if the following series converges or diverges. If it converges determine its sum.

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

Solution

This is actually one of the few series in which we are able to determine a formula for the general term in the sequence of partial fractions. However, in this section we are more interested in the general idea of convergence and divergence and so we'll put off discussing the process for finding the formula until the next section.

The general formula for the partial sums is,

$$s_n = \sum_{i=2}^{n} \frac{1}{i^2-1} = \frac{3}{4} - \frac{1}{2n} - \frac{1}{2(n+1)}$$

and in this case we have,

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty}\left(\frac{3}{4} - \frac{1}{2n} - \frac{1}{2(n+1)}\right) = \frac{3}{4}$$

The sequence of partial sums converges and so the series converges also and its value is,

$$\sum_{n=2}^{\infty} \frac{1}{n^2-1} = \frac{3}{4}$$

Example 3 Determine if the following series converges or diverges. If it converges determine its sum.

$$\sum_{n=0}^{\infty} (-1)^n$$

Solution

In this case we really don't need a general formula for the partial sums to determine the convergence of this series. Let's just write down the first few partial sums.

$$\begin{aligned} &s_0 = 1 \\ &s_1 = 1-1 = 0 \\ &s_2 = 1-1+1 = 1 \\ &s_3 = 1-1+1-1 = 0 \\ &etc. \end{aligned}$$

So, it looks like the sequence of partial sums is,

$$\{s_n\}_{n=0}^{\infty} = \{1,0,1,0,1,0,1,0,1,\ldots\}$$

and this sequence diverges since $\lim\limits_{n\to\infty} s_n$ doesn't exist. Therefore, the series also diverges.

Example 4 Determine if the following series converges or diverges. If it converges determine its sum.

$$\sum_{n=1}^{\infty} \frac{1}{3^{n-1}}$$

Solution

Here is the general formula for the partial sums for this series.

$$s_n = \sum_{i=1}^{n} \frac{1}{3^{i-1}} = \frac{3}{2}\left(1 - \frac{1}{3^n}\right)$$

Again, do not worry about knowing this formula. This is not something that you'll ever be asked to know in my class.

In this case the limit of the sequence of partial sums is,

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{3}{2}\left(1 - \frac{1}{3^n}\right) = \frac{3}{2}$$

The sequence of partial sums is convergent and so the series will also be convergent. The value of the series is,

$$\sum_{n=1}^{\infty} \frac{1}{3^{n-1}} = \frac{3}{2}$$

As we already noted, do not get excited about determining the general formula for the sequence of partial sums. There is only going to be one type of series where you will need to determine this formula and the process in that case isn't too bad. In fact, you already know how to do most of the work in the process as you'll see in the next section.

So, we've determined the convergence of four series now. Two of the series converged and two diverged. Let's go back and examine the series terms for each of these. For each of the series let's take the limit as n goes to infinity of the series terms (not the partial sums!!).

$\lim_{n\to\infty} n = \infty$	this series diverged
$\lim_{n\to\infty} \frac{1}{n^2 - 1} = 0$	this series converged
$\lim_{n\to\infty} (-1)^n$ doesn't exist	this series diverged
$\lim_{n\to\infty} \frac{1}{3^{n-1}} = 0$	this series converged

Notice that for the two series that converged the series term itself was zero in the limit. This will always be true for convergent series and leads to the following theorem.

Theorem

If $\sum a_n$ converges then $\lim_{n\to\infty} a_n = 0$.

Proof

First let's suppose that the series starts at $n = 1$. If it doesn't then we can modify things as appropriate below. Then the partial sums are,

$$s_{n-1} = \sum_{i=1}^{n-1} a_i = a_1 + a_2 + a_3 + a_4 + \cdots + a_{n-1} \qquad s_n = \sum_{i=1}^{n} a_i = a_1 + a_2 + a_3 + a_4 + \cdots + a_{n-1} + a_n$$

Next, we can use these two partial sums to write,

$$a_n = s_n - s_{n-1}$$

Now because we know that $\sum a_n$ is convergent we also know that the sequence $\left\{s_n\right\}_{n=1}^{\infty}$ is also convergent and that $\lim\limits_{n\to\infty} s_n = s$ for some finite value *s*. However, since $n-1 \to \infty$ as $n \to \infty$ we also have $\lim\limits_{n\to\infty} s_{n-1} = s$.

We now have,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}\left(s_n - s_{n-1}\right) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = s - s = 0$$

■

Be careful to not misuse this theorem! This theorem gives us a requirement for convergence but not a guarantee of convergence. In other words, the converse is NOT true. If $\lim\limits_{n\to\infty} a_n = 0$ the series may actually diverge! Consider the following two series.

$$\sum_{n=1}^{\infty} \frac{1}{n} \qquad \sum_{n=1}^{\infty} \frac{1}{n^2}$$

In both cases the series terms are zero in the limit as *n* goes to infinity, yet only the second series converges. The first series diverges. It will be a couple of sections before we can prove this, so at this point please believe this and know that you'll be able to prove the convergence of these two series in a couple of sections.

Again, as noted above, all this theorem does is give us a requirement for a series to converge. In order for a series to converge the series terms must go to zero in the limit. If the series terms do not go to zero in the limit then there is no way the series can converge since this would violate the theorem.

This leads us to the first of many tests for the convergence/divergence of a series that we'll be seeing in this chapter.

Divergence Test

If $\lim\limits_{n\to\infty} a_n \neq 0$ then $\sum a_n$ will diverge.

Again, do NOT misuse this test. This test only says that a series is guaranteed to diverge if the series terms don't go to zero in the limit. If the series terms do happen to go to zero the series may or may not converge! Again, recall the following two series,

$$\sum_{n=1}^{\infty} \frac{1}{n} \qquad \text{diverges}$$

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \qquad \text{converges}$$

One of the more common mistakes that students make when they first get into series is to assume that if $\lim\limits_{n \to \infty} a_n = 0$ then $\sum a_n$ will converge. There is just no way to guarantee this so be careful!

Let's take a quick look at an example of how this test can be used.

Example 5 Determine if the following series is convergent or divergent.

$$\sum_{n=0}^{\infty} \frac{4n^2 - n^3}{10 + 2n^3}$$

Solution

With almost every series we'll be looking at in this chapter the first thing that we should do is take a look at the series terms and see if they go to zero or not. If it's clear that the terms don't go to zero use the Divergence Test and be done with the problem.

That's what we'll do here.

$$\lim_{n \to \infty} \frac{4n^2 - n^3}{10 + 2n^3} = -\frac{1}{2} \neq 0$$

The limit of the series terms isn't zero and so by the Divergence Test the series diverges.

The divergence test is the first test of many tests that we will be looking at over the course of the next several sections. You will need to keep track of all these tests, the conditions under which they can be used and their conclusions all in one place so you can quickly refer back to them as you need to.

Next we should briefly revisit arithmetic of series and convergence/divergence. As we saw in the previous section if $\sum a_n$ and $\sum b_n$ are both convergent series then so are $\sum ca_n$ and $\sum_{n=k}^{\infty} (a_n \pm b_n)$. Furthermore, these series will have the following sums or values.

$$\sum ca_n = c\sum a_n \qquad \sum_{n=k}^{\infty} (a_n \pm b_n) = \sum_{n=k}^{\infty} a_n \pm \sum_{n=k}^{\infty} b_n$$

We'll see an example of this in the next section after we get a few more examples under our belt. At this point just remember that a sum of convergent series is convergent and multiplying a convergent series by a number will not change its convergence.

We need to be a little careful with these facts when it comes to divergent series. In the first case if $\sum a_n$ is divergent then $\sum ca_n$ will also be divergent (provided c isn't zero of course) since multiplying a series that is infinite in value or doesn't have a value by a finite value (*i.e.* c) won't change the fact that the series has an infinite or no value. However, it is possible to have both $\sum a_n$ and $\sum b_n$ be divergent series and yet have $\sum_{n=k}^{\infty}\left(a_n \pm b_n\right)$ be a convergent series.

Now, since the main topic of this section is the convergence of a series we should mention a stronger type of convergence. A series $\sum a_n$ is said to **converge absolutely** if $\sum \left|a_n\right|$ also converges. Absolute convergence is *stronger* than convergence in the sense that a series that is absolutely convergent will also be convergent, but a series that is convergent may or may not be absolutely convergent.

In fact if $\sum a_n$ converges and $\sum \left|a_n\right|$ diverges the series $\sum a_n$ is called **conditionally convergent**.

At this point we don't really have the tools at hand to properly investigate this topic in detail nor do we have the tools in hand to determine if a series is absolutely convergent or not. So we'll not say anything more about this subject for a while. When we finally have the tools in hand to discuss this topic in more detail we will revisit it. Until then don't worry about it. The idea is mentioned here only because we were already discussing convergence in this section and it ties into the last topic that we want to discuss in this section.

In the previous section after we'd introduced the idea of an infinite series we commented on the fact that we shouldn't think of an infinite series as an infinite sum despite the fact that the notation we use for infinite series seems to imply that it is an infinite sum. It's now time to briefly discuss this.

First, we need to introduce the idea of a **rearrangement**. A rearrangement of a series is exactly what it might sound like, it is the same series with the terms rearranged into a different order.

For example, consider the following infinite series.

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + \cdots$$

A rearrangement of this series is,

$$\sum_{n=1}^{\infty} a_n = a_2 + a_1 + a_3 + a_{14} + a_5 + a_9 + a_4 + \cdots$$

The issue we need to discuss here is that for some series each of these arrangements of terms can have different values despite the fact that they are using exactly the same terms.

Here is an example of this. It can be shown that,

$$\sum_{n=1}^{\infty} \frac{\left(-1\right)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots = \ln 2 \qquad (1)$$

Since this series converges we know that if we multiply it by a constant *c* its value will also be multiplied by *c*. So, let's multiply this by $\frac{1}{2}$ to get,

$$\frac{1}{2}-\frac{1}{4}+\frac{1}{6}-\frac{1}{8}+\frac{1}{10}-\frac{1}{12}+\frac{1}{14}-\frac{1}{16}+\cdots=\frac{1}{2}\ln 2 \tag{2}$$

Now, let's add in a zero between each term as follows.

$$0+\frac{1}{2}+0-\frac{1}{4}+0+\frac{1}{6}+0-\frac{1}{8}+0+\frac{1}{10}+0-\frac{1}{12}+0+\cdots=\frac{1}{2}\ln 2 \tag{3}$$

Note that this won't change the value of the series because the partial sums for this series will be the partial sums for the (2) except that each term will be repeated. Repeating terms in a series will not affect its limit however and so both (2) and (3) will be the same.

We know that if two series converge we can add them by adding term by term and so add (1) and (3) to get,

$$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\cdots=\frac{3}{2}\ln 2 \tag{4}$$

Now, notice that the terms of (4) are simply the terms of (1) rearranged so that each negative term comes after two positive terms. The values however are definitely different despite the fact that the terms are the same.

Note as well that this is not one of those "tricks" that you see occasionally where you get a contradictory result because of a hard to spot math/logic error. This is a very real result and we've not made any logic mistakes/errors.

Here is a nice set of facts that govern this idea of when a rearrangement will lead to a different value of a series.

Facts

Given the series $\sum a_n$,

1. If $\sum a_n$ is absolutely convergent and its value is *s* then any rearrangement of $\sum a_n$ will also have a value of *s*.

2. If $\sum a_n$ is conditionally convergent and *r* is any real number then there is a rearrangement of $\sum a_n$ whose value will be *r*.

Again, we do not have the tools in hand yet to determine if a series is absolutely convergent and so don't worry about this at this point. This is here just to make sure that you understand that we have to be very careful in thinking of an infinite series as an infinite sum. There are times when we can (*i.e.* the series is absolutely convergent) and there are times when we can't (*i.e.* the series is conditionally convergent).

As a final note, the fact above tells us that the series,

$$\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}$$

must be conditionally convergent since two rearrangements gave two separate values of this series. Eventually it will be very simple to show that this series is conditionally convergent.

Series – Special Series

In this section we are going to take a brief look at three special series. Actually, special may not be the correct term. All three have been named which makes them special in some way, however the main reason that we're going to look at two of them in this section is that they are the only types of series that we'll be looking at for which we will be able to get actual values for the series. The third type is divergent and so won't have a value to worry about.

In general, determining the value of a series is very difficult and outside of these two kinds of series that we'll look at in this section we will not be determining the value of series in this chapter.

So, let's get started.

Geometric Series
A geometric series is any series that can be written in the form,

$$\sum_{n=1}^{\infty} ar^{n-1}$$

or, with an index shift the geometric series will often be written as,

$$\sum_{n=0}^{\infty} ar^{n}$$

These are identical series and will have identical values, provided they converge of course.

If we start with the first form it can be shown that the partial sums are,

$$s_n = \frac{a\left(1-r^n\right)}{1-r} = \frac{a}{1-r} - \frac{ar^n}{1-r}$$

The series will converge provided the partial sums form a convergent sequence, so let's take the limit of the partial sums.

$$\begin{aligned} \lim_{n\to\infty} s_n &= \lim_{n\to\infty}\left(\frac{a}{1-r} - \frac{ar^n}{1-r}\right) \\ &= \lim_{n\to\infty}\frac{a}{1-r} - \lim_{n\to\infty}\frac{ar^n}{1-r} \\ &= \frac{a}{1-r} - \frac{a}{1-r}\lim_{n\to\infty} r^n \end{aligned}$$

Now, from Theorem 3 from the Sequences section we know that the limit above will exist and be finite provided $-1 < r \le 1$. However, note that we can't let $r = 1$ since this will give division by zero. Therefore, this will exist and be finite provided $-1 < r < 1$ and in this case the limit is zero and so we get,

$$\lim_{n\to\infty} s_n = \frac{a}{1-r}$$

Therefore, a geometric series will converge if $-1 < r < 1$, which is usually written $|r| < 1$, its value is,

$$\sum_{n=1}^{\infty} ar^{n-1} = \sum_{n=0}^{\infty} ar^{n} = \frac{a}{1-r}$$

Note that in using this formula we'll need to make sure that we are in the correct form. In other words, if the series starts at $n = 0$ then the exponent on the r must be n. Likewise if the series starts at $n = 1$ then the exponent on the r must be $n - 1$.

Example 1 Determine if the following series converge or diverge. If they converge give the value of the series.

(a) $\displaystyle\sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1}$ [Solution]

(b) $\displaystyle\sum_{n=0}^{\infty} \frac{(-4)^{3n}}{5^{n-1}}$ [Solution]

Solution

(a) $\displaystyle\sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1}$

This series doesn't really look like a geometric series. However, notice that both parts of the series term are numbers raised to a power. This means that it can be put into the form of a geometric series. We will just need to decide which form is the correct form. Since the series starts at $n = 1$ we will want the exponents on the numbers to be $n - 1$.

It will be fairly easy to get this into the correct form. Let's first rewrite things slightly. One of the n's in the exponent has a negative in front of it and that can't be there in the geometric form. So, let's first get rid of that.

$$\sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1} = \sum_{n=1}^{\infty} 9^{-(n-2)} 4^{n+1} = \sum_{n=1}^{\infty} \frac{4^{n+1}}{9^{n-2}}$$

Now let's get the correct exponent on each of the numbers. This can be done using simple exponent properties.

$$\sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1} = \sum_{n=1}^{\infty} \frac{4^{n+1}}{9^{n-2}} = \sum_{n=1}^{\infty} \frac{4^{n-1} 4^{2}}{9^{n-1} 9^{-1}}$$

Now, rewrite the term a little.

$$\sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1} = \sum_{n=1}^{\infty} 16(9) \frac{4^{n-1}}{9^{n-1}} = \sum_{n=1}^{\infty} 144 \left(\frac{4}{9}\right)^{n-1}$$

So, this is a geometric series with $a = 144$ and $r = \frac{4}{9} < 1$. Therefore, since $|r| < 1$ we know the series will converge and its value will be,

$$\sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1} = \frac{144}{1-\frac{4}{9}} = \frac{9}{5}(144) = \frac{1296}{5}$$

[Return to Problems]

(b) $\displaystyle\sum_{n=0}^{\infty} \frac{(-4)^{3n}}{5^{n-1}}$

Again, this doesn't look like a geometric series, but it can be put into the correct form. In this case the series starts at $n=0$ so we'll need the exponents to be n on the terms. Note that this means we're going to need to rewrite the exponent on the numerator a little

$$\sum_{n=0}^{\infty} \frac{(-4)^{3n}}{5^{n-1}} = \sum_{n=0}^{\infty} \frac{\left((-4)^3\right)^n}{5^n 5^{-1}} = \sum_{n=0}^{\infty} 5\frac{(-64)^n}{5^n} = \sum_{n=0}^{\infty} 5\left(\frac{-64}{5}\right)^n$$

So, we've got it into the correct form and we can see that $a=5$ and $r=-\frac{64}{5}$. Also note that $|r| \geq 1$ and so this series diverges.

[Return to Problems]

Back in the Series – Basics section we talked about stripping out terms from a series, but didn't really provide any examples of how this idea could be used in practice. We can now do some examples.

Example 2 Use the results from the previous example to determine the value of the following series.

(a) $\displaystyle\sum_{n=0}^{\infty} 9^{-n+2} 4^{n+1}$ [Solution]

(b) $\displaystyle\sum_{n=3}^{\infty} 9^{-n+2} 4^{n+1}$ [Solution]

Solution

(a) $\displaystyle\sum_{n=0}^{\infty} 9^{-n+2} 4^{n+1}$

In this case we could just acknowledge that this is a geometric series that starts at $n=0$ and so we could put it into the correct form and be done with it. However, this does provide us with a nice example of how to use the idea of stripping out terms to our advantage.

Let's notice that if we strip out the first term from this series we arrive at,

$$\sum_{n=0}^{\infty} 9^{-n+2} 4^{n+1} = 9^2 4^1 + \sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1} = 324 + \sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1}$$

From the previous example we know the value of the new series that arises here and so the value of the series in this example is,

$$\sum_{n=0}^{\infty} 9^{-n+2} 4^{n+1} = 324 + \frac{1296}{5} = \frac{2916}{5}$$

[Return to Problems]

(b) $\sum_{n=3}^{\infty} 9^{-n+2} 4^{n+1}$

In this case we can't strip out terms from the given series to arrive at the series used in the previous example. However, we can start with the series used in the previous example and strip terms out of it to get the series in this example. So, let's do that. We will strip out the first two terms from the series we looked at in the previous example.

$$\sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1} = 9^1 4^2 + 9^0 4^3 + \sum_{n=3}^{\infty} 9^{-n+2} 4^{n+1} = 208 + \sum_{n=3}^{\infty} 9^{-n+2} 4^{n+1}$$

We can now use the value of the series from the previous example to get the value of this series.

$$\sum_{n=3}^{\infty} 9^{-n+2} 4^{n+1} = \sum_{n=1}^{\infty} 9^{-n+2} 4^{n+1} - 208 = \frac{1296}{5} - 208 = \frac{256}{5}$$

[Return to Problems]

Notice that we didn't discuss the convergence of either of the series in the above example. Here's why. Consider the following series written in two separate ways (*i.e.* we stripped out a couple of terms from it).

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \sum_{n=3}^{\infty} a_n$$

Let's suppose that we know $\sum_{n=3}^{\infty} a_n$ is a convergent series. This means that it's got a finite value and adding three finite terms onto this will not change that fact. So the value of $\sum_{n=0}^{\infty} a_n$ is also finite and so is convergent.

Likewise, suppose that $\sum_{n=0}^{\infty} a_n$ is convergent. In this case if we subtract three finite values from this value we will remain finite and arrive at the value of $\sum_{n=3}^{\infty} a_n$. This is now a finite value and so this series will also be convergent.

In other words, if we have two series and they differ only by the presence, or absence, of a finite number of finite terms they will either both be convergent or they will both be divergent. The difference of a few terms one way or the other will not change the convergence of a series. This is an important idea and we will use it several times in the following sections to simplify some of the tests that we'll be looking at.

Telescoping Series
It's now time to look at the second of the three series in this section. In this portion we are going to look at a series that is called a telescoping series. The name in this case comes from what happens with the partial sums and is best shown in an example.

Example 3 Determine if the following series converges or diverges. If it converges find its value.

$$\sum_{n=0}^{\infty} \frac{1}{n^2+3n+2}$$

Solution
We first need the partial sums for this series.

$$s_n = \sum_{i=0}^{n} \frac{1}{i^2+3i+2}$$

Now, let's notice that we can use partial fractions on the series term to get,

$$\frac{1}{i^2+3i+2} = \frac{1}{(i+2)(i+1)} = \frac{1}{i+1} - \frac{1}{i+2}$$

I'll leave the details of the partial fractions to you. By now you should be fairly adept at this since we spent a fair amount of time doing partial fractions back in the Integration Techniques chapter. If you need a refresher you should go back and review that section.

So, what does this do for us? Well, let's start writing out the terms of the general partial sum for this series using the partial fraction form.

$$\begin{aligned} s_n &= \sum_{i=0}^{n} \left(\frac{1}{i+1} - \frac{1}{i+2} \right) \\ &= \left(\frac{1}{1} - \cancel{\frac{1}{2}} \right) + \left(\cancel{\frac{1}{2}} - \cancel{\frac{1}{3}} \right) + \left(\cancel{\frac{1}{3}} - \cancel{\frac{1}{4}} \right) + \cdots + \left(\cancel{\frac{1}{n}} - \cancel{\frac{1}{n+1}} \right) + \left(\cancel{\frac{1}{n+1}} - \frac{1}{n+2} \right) \\ &= 1 - \frac{1}{n+2} \end{aligned}$$

Notice that every term except the first and last term canceled out. This is the origin of the name **telescoping series**.

This also means that we can determine the convergence of this series by taking the limit of the partial sums.

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \left(1 - \frac{1}{n+2} \right) = 1$$

The sequence of partial sums is convergent and so the series is convergent and has a value of

$$\sum_{n=0}^{\infty} \frac{1}{n^2+3n+2} = 1$$

In telescoping series be careful to not assume that successive terms will be the ones that cancel. Consider the following example.

Example 4 Determine if the following series converges or diverges. If it converges find its value.

$$\sum_{n=1}^{\infty} \frac{1}{n^2 + 4n + 3}$$

Solution

As with the last example we'll leave the partial fractions details to you to verify. The partial sums are,

$$\begin{aligned} s_n &= \sum_{i=1}^{n} \left(\frac{\frac{1}{2}}{i+1} - \frac{\frac{1}{2}}{i+3} \right) = \frac{1}{2} \sum_{i=1}^{n} \left(\frac{1}{i+1} - \frac{1}{i+3} \right) \\ &= \frac{1}{2} \left[\left(\frac{1}{2} - \cancel{\frac{1}{4}} \right) + \left(\frac{1}{3} - \cancel{\frac{1}{5}} \right) + \left(\cancel{\frac{1}{4}} - \cancel{\frac{1}{6}} \right) + \cdots + \left(\cancel{\frac{1}{n}} - \frac{1}{n+2} \right) + \left(\cancel{\frac{1}{n+1}} - \frac{1}{n+3} \right) \right] \\ &= \frac{1}{2} \left[\frac{1}{2} + \frac{1}{3} - \frac{1}{n+2} - \frac{1}{n+3} \right] \end{aligned}$$

In this case instead of successive terms canceling a term will cancel with a term that is farther down the list. The end result this time is two initial and two final terms are left. Notice as well that in order to help with the work a little we factored the $\frac{1}{2}$ out of the series.

The limit of the partial sums is,

$$\lim_{n \to \infty} s_n = \lim_{n \to \infty} \frac{1}{2} \left(\frac{5}{6} - \frac{1}{n+2} - \frac{1}{n+3} \right) = \frac{5}{12}$$

So, this series is convergent (because the partial sums form a convergent sequence) and its value is,

$$\sum_{n=1}^{\infty} \frac{1}{n^2 + 4n + 3} = \frac{5}{12}$$

Note that it's not always obvious if a series is telescoping or not until you try to get the partial sums and then see if they are in fact telescoping. There is no test that will tell us that we've got a telescoping series right off the bat. Also note that just because you can do partial fractions on a series term does not mean that the series will be a telescoping series. The following series, for example, is not a telescoping series despite the fact that we can partial fraction the series terms.

$$\sum_{n=1}^{\infty} \frac{3 + 2n}{n^2 + 3n + 2} = \sum_{n=1}^{\infty} \left(\frac{1}{n+1} + \frac{1}{n+2} \right)$$

In order for a series to be a telescoping series we must get terms to cancel and all of these terms are positive and so none will cancel.

Next, we need to go back and address an issue that was first raised in the previous section. In that section we stated that the sum or difference of convergent series was also convergent and that the presence of a multiplicative constant would not affect the convergence of a series. Now that we have a few more series in hand let's work a quick example showing that.

Example 5 Determine the value of the following series.

$$\sum_{n=1}^{\infty}\left(\frac{4}{n^2+4n+3}-9^{-n+2}4^{n+1}\right)$$

Solution

To get the value of this series all we need to do is rewrite it and then use the previous results.

$$\begin{aligned}\sum_{n=1}^{\infty}\left(\frac{4}{n^2+4n+3}-9^{-n+2}4^{n+1}\right) &= \sum_{n=1}^{\infty}\frac{4}{n^2+4n+3}-\sum_{n=1}^{\infty}9^{-n+2}4^{n+1}\\ &= 4\sum_{n=1}^{\infty}\frac{1}{n^2+4n+3}-\sum_{n=1}^{\infty}9^{-n+2}4^{n+1}\\ &= 4\left(\frac{5}{12}\right)-\frac{1296}{5}\\ &= -\frac{3863}{15}\end{aligned}$$

We didn't discuss the convergence of this series because it was the sum of two convergent series and that guaranteed that the original series would also be convergent.

Harmonic Series

This is the third and final series that we're going to look at in this section. Here is the harmonic series.

$$\sum_{n=1}^{\infty}\frac{1}{n}$$

The harmonic series is divergent and we'll need to wait until the next section to show that. This series is here because it's got a name and so I wanted to put it here with the other two named series that we looked at in this section. We're also going to use the harmonic series to illustrate a couple of ideas about divergent series that we've already discussed for convergent series. We'll do that with the following example.

Example 6 Show that each of the following series are divergent.

(a) $\displaystyle\sum_{n=1}^{\infty}\frac{5}{n}$

(b) $\displaystyle\sum_{n=4}^{\infty}\frac{1}{n}$

Solution

(a) $\displaystyle\sum_{n=1}^{\infty}\frac{5}{n}$

To see that this series is divergent all we need to do is use the fact that we can factor a constant out of a series as follows,

$$\sum_{n=1}^{\infty}\frac{5}{n}=5\sum_{n=1}^{\infty}\frac{1}{n}$$

Now, $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent and so five times this will still not be a finite number and so the series has to be divergent. In other words, if we multiply a divergent series by a constant it will still be divergent.

(b) $\sum_{n=4}^{\infty} \frac{1}{n}$

In this case we'll start with the harmonic series and strip out the first three terms.

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \sum_{n=4}^{\infty} \frac{1}{n} \qquad \Rightarrow \qquad \sum_{n=4}^{\infty} \frac{1}{n} = \left(\sum_{n=1}^{\infty} \frac{1}{n} \right) - \frac{11}{6}$$

In this case we are subtracting a finite number from a divergent series. This subtraction will not change the divergence of the series. We will either have infinity minus a finite number, which is still infinity, or a series with no value minus a finite number, which will still have no value.

Therefore, this series is divergent.

Just like with convergent series, adding/subtracting a finite number from a divergent series is not going to change the divergence of the series.

So, some general rules about the convergence/divergence of a series are now in order. Multiplying a series by a constant will not change the convergence/divergence of the series and adding or subtracting a constant from a series will not change the convergence/divergence of the series. These are nice ideas to keep in mind.

Integral Test

The last topic that we discussed in the previous section was the harmonic series. In that discussion we stated that the harmonic series was a divergent series. It is now time to prove that statement. This proof will also get us started on the way to our next test for convergence that we'll be looking at.

So, we will be trying to prove that the harmonic series,

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges.

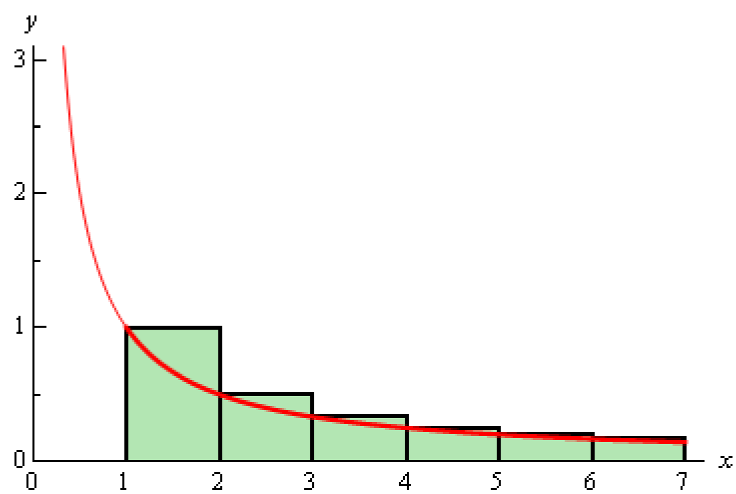
We'll start this off by looking at an apparently unrelated problem. Let's start off by asking what the area under $f(x) = \frac{1}{x}$ on the interval $[1, \infty)$. From the section on Improper Integrals we know that this is,

$$\int_1^{\infty} \frac{1}{x}\, dx = \infty$$

and so we called this integral divergent (yes, that's the same term we're using here with series….).

So, just how does that help us to prove that the harmonic series diverges? Well, recall that we can always estimate the area by breaking up the interval into segments and then sketching in rectangles and using the sum of the area all of the rectangles as an estimate of the actual area. Let's do that for this problem as well and see what we get.

We will break up the interval into subintervals of width 1 and we'll take the function value at the left endpoint as the height of the rectangle. The image below shows the first few rectangles for this area.



So, the area under the curve is approximately,

$$
\begin{aligned}
A &\approx \left(\frac{1}{1}\right)(1)+\left(\frac{1}{2}\right)(1)+\left(\frac{1}{3}\right)(1)+\left(\frac{1}{4}\right)(1)+\left(\frac{1}{5}\right)(1)+\cdots \\
&=\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots \\
&=\sum_{n=1}^{\infty}\frac{1}{n}
\end{aligned}
$$

Now note a couple of things about this approximation. First, each of the rectangles overestimates the actual area and secondly the formula for the area is exactly the harmonic series!

Putting these two facts together gives the following,

$$A \approx \sum_{n=1}^{\infty}\frac{1}{n} > \int_{1}^{\infty}\frac{1}{x}\,dx = \infty$$

Notice that this tells us that we must have,

$$\sum_{n=1}^{\infty}\frac{1}{n} > \infty \qquad\qquad \Rightarrow \qquad\qquad \sum_{n=1}^{\infty}\frac{1}{n} = \infty$$

Since we can't really be larger than infinity the harmonic series must also be infinite in value. In other words, the harmonic series is in fact divergent.

So, we've managed to relate a series to an improper integral that we could compute and it turns out that the improper integral and the series have exactly the same convergence.

Let's see if this will also be true for a series that converges. When discussing the Divergence Test we made the claim that

$$\sum_{n=1}^{\infty}\frac{1}{n^2}$$

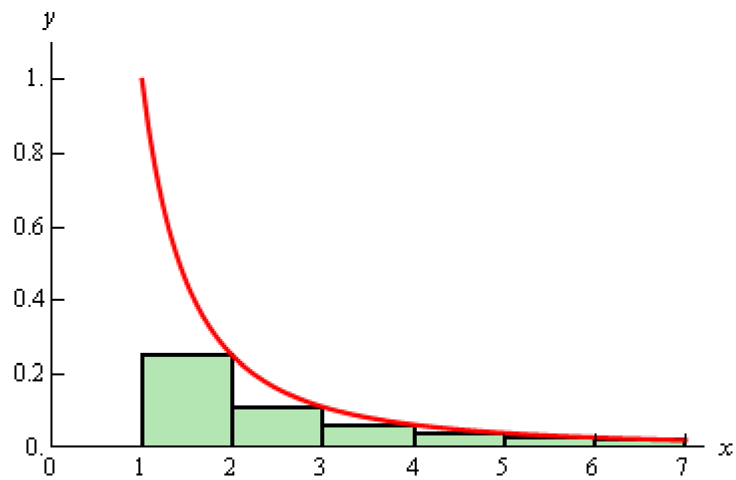
converges. Let's see if we can do something similar to the above process to prove this.

We will try to relate this to the area under $f(x)=\frac{1}{x^2}$ is on the interval $[1,\infty)$. Again, from the Improper Integral section we know that,

$$\int_{1}^{\infty}\frac{1}{x^2}\,dx = 1$$

and so this integral converges.

We will once again try to estimate the area under this curve. We will do this in an almost identical manner as the previous part with the exception that instead of using the left end points for the height of our rectangles we will use the right end points. Here is a sketch of this case,



In this case the area estimation is,

$$
\begin{aligned}
A &\approx \left(\frac{1}{2^2}\right)(1)+\left(\frac{1}{3^2}\right)(1)+\left(\frac{1}{4^2}\right)(1)+\left(\frac{1}{5^2}\right)(1)+\cdots \\
&= \frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\frac{1}{5^2}+\cdots
\end{aligned}
$$

This time, unlike the first case, the area will be an underestimation of the actual area and the estimation is not quite the series that we are working with. Notice however that the only difference is that we're missing the first term. This means we can do the following,

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{1}{1^2}+\underbrace{\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\frac{1}{5^2}+\cdots}_{\text{Area Estimation}}<1+\int_1^{\infty}\frac{1}{x^2}\,dx=1+1=2$$

Or, putting all this together we see that,

$$\sum_{n=1}^{\infty}\frac{1}{n^2}<2$$

With the harmonic series this was all that we needed to say that the series was divergent. With this series however, this isn't quite enough. For instance $-\infty<2$ and if the series did have a value of $-\infty$ then it would be divergent (when we want convergent). So, let's do a little more work.

First, let's notice that all the series terms are positive (that's important) and that the partial sums are,

$$s_n=\sum_{i=1}^{n}\frac{1}{i^2}$$

Because the terms are all positive we know that the partial sums must be an increasing sequence. In other words,

$$s_n=\sum_{i=1}^{n}\frac{1}{i^2}<\sum_{i=1}^{n+1}\frac{1}{i^2}=s_{n+1}$$

In s_{n+1} we are adding a single positive term onto s_n and so must get larger. Therefore, the partial sums form an increasing (and hence monotonic) sequence.

Also note that, since the terms are all positive, we can say,

$$s_n = \sum_{i=1}^{n} \frac{1}{i^2} < \sum_{n=1}^{\infty} \frac{1}{n^2} < 2 \qquad \Rightarrow \qquad s_n < 2$$

and so the sequence of partial sums is a bounded sequence.

In the second section on Sequences we gave a theorem that stated that a bounded and monotonic sequence was guaranteed to be convergent. This means that the sequence of partial sums is a convergent sequence. So, who cares right? Well recall that this means that the series must then also be convergent!

So, once again we were able to relate a series to an improper integral (that we could compute) and the series and the integral had the same convergence.

We went through a fair amount of work in both of these examples to determine the convergence of the two series. Luckily for us we don't need to do all this work every time. The ideas in these two examples can be summarized in the following test.

Integral Test

Suppose that $f(x)$ is a continuous, positive and decreasing function on the interval $[k, \infty)$ and that $f(n) = a_n$ then,

1. If $\int_k^{\infty} f(x)\,dx$ is convergent so is $\sum_{n=k}^{\infty} a_n$.
2. If $\int_k^{\infty} f(x)\,dx$ is divergent so is $\sum_{n=k}^{\infty} a_n$.

A formal proof of this test can be found at the end of this section.

There are a couple of things to note about the integral test. First, the lower limit on the improper integral must be the same value that starts the series.

Second, the function does not actually need to be decreasing and positive everywhere in the interval. All that's really required is that eventually the function is decreasing and positive. In other words, it is okay if the function (and hence series terms) increases or is negative for a while, but eventually the function (series terms) must decrease and be positive for all terms. To see why this is true let's suppose that the series terms increase and or are negative in the range $k \le n \le N$ and then decrease and are positive for $n \ge N+1$. In this case the series can be written as,

$$\sum_{n=k}^{\infty} a_n = \sum_{n=k}^{N} a_n + \sum_{n=N+1}^{\infty} a_n$$

Now, the first series is nothing more than a finite sum (no matter how large N is) of finite terms and so will be finite. So the original series will be convergent/divergent only if the second infinite series on the right is convergent/divergent and the test can be done on the second series as it satisfies the conditions of the test.

A similar argument can be made using the improper integral as well.

The requirement in the test that the function/series be decreasing and positive everywhere in the range is required for the proof. In practice however, we only need to make sure that the function/series is eventually a decreasing and positive function/series. Also note that when computing the integral in the test we don't actually need to strip out the increasing/negative portion since the presence of a small range on which the function is increasing/negative will not change the integral from convergent to divergent or from divergent to convergent.

There is one more very important point that must be made about this test. This test does NOT give the value of a series. It will only give the convergence/divergence of the series. That's it. No value. We can use the above series as a perfect example of this. All that the test gave us was that,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} < 2$$

So, we got an upper bound on the value of the series, but not an actual value for the series. In fact, from this point on we will not be asking for the value of a series we will only be asking whether a series converges or diverges. In a later section we look at estimating values of series, but even in that section still won't actually be getting values of series.

Just for the sake of completeness the value of this series is known.

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} = 1.644934\ldots < 2$$

Let's work a couple of examples.

Example 1 Determine if the following series is convergent or divergent.

$$\sum_{n=2}^{\infty} \frac{1}{n \ln n}$$

Solution

In this case the function we'll use is,

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

This function is clearly positive and if we make x larger the denominator will get larger and so the function is also decreasing. Therefore, all we need to do is determine the convergence of the following integral.

$$\begin{aligned} \int_2^{\infty} \frac{1}{x \ln x}\, dx &= \lim_{t \to \infty} \int_2^{t} \frac{1}{x \ln x}\, dx \qquad\qquad u = \ln x \\ &= \lim_{t \to \infty} \left(\ln(\ln x) \right) \Big|_2^t \\ &= \lim_{t \to \infty} \left(\ln(\ln t) - \ln(\ln 2) \right) \\ &= \infty \end{aligned}$$

The integral is divergent and so the series is also divergent by the Integral Test.

Example 2 Determine if the following series is convergent or divergent.

$$\sum_{n=0}^{\infty} n\mathbf{e}^{-n^2}$$

Solution

The function that we'll use in this example is,

$$f(x) = x\mathbf{e}^{-x^2}$$

This function is always positive on the interval that we're looking at. Now we need to check that the function is decreasing. It is not clear that this function will always be decreasing on the interval given. We can use our Calculus I knowledge to help us however. The derivative of this function is,

$$f'(x) = \mathbf{e}^{-x^2}\left(1-2x^2\right)$$

This function has two critical points (which will tell us where the derivative changes sign) at $x = \pm\frac{1}{\sqrt{2}}$. Since we are starting at $n = 0$ we can ignore the negative critical point. Picking a couple of test points we can see that the function is increasing on the interval $\left[0, \frac{1}{\sqrt{2}}\right]$ and it is decreasing on $\left[\frac{1}{\sqrt{2}}, \infty\right)$. Therefore, eventually the function will be decreasing and that's all that's required for us to use the Integral Test.

$$\begin{aligned} \int_0^{\infty} x\mathbf{e}^{-x^2}\,dx &= \lim_{t\to\infty}\int_0^t x\mathbf{e}^{-x^2}\,dx \qquad\qquad u = -x^2 \\ &= \lim_{t\to\infty}\left(-\frac{1}{2}\mathbf{e}^{-x^2}\right)\Bigg|_0^t \\ &= \lim_{t\to\infty}\left(\frac{1}{2}-\frac{1}{2}\mathbf{e}^{-t^2}\right) = \frac{1}{2} \end{aligned}$$

The integral is convergent and so the series must also be convergent by the Integral Test.

We can use the Integral Test to get the following fact/test for some series.

Fact (The *p* –series Test)

If $k > 0$ then $\displaystyle\sum_{n=k}^{\infty}\frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \le 1$.

Sometimes the series in this fact are called *p*-series and so this fact is sometimes called the *p*-series test. This fact follows directly from the Integral Test and a similar fact we saw in the Improper Integral section. This fact says that the integral,

$$\int_k^{\infty}\frac{1}{x^p}\,dx$$

converges if $p > 1$ and diverges if $p \le 1$.

Using the *p*-series test makes it very easy to determine the convergence of some series.

Example 3 Determine if the following series are convergent or divergent.

(a) $\displaystyle\sum_{n=4}^{\infty}\frac{1}{n^7}$

(b) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}$

Solution

(a) In this case $p = 7 > 1$ and so by this fact the series is convergent.

(b) For this series $p = \frac{1}{2} \le 1$ and so the series is divergent by the fact.

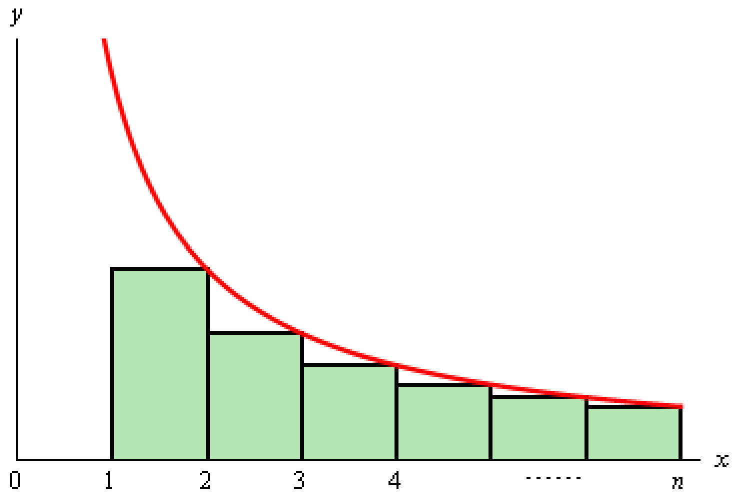
The last thing that we'll do in this section is give a quick proof of the Integral Test. We've essentially done the proof already at the beginning of the section when we were introducing the Integral Test, but let's go through it formally for a general function.

Proof of Integral Test

First, for the sake of the proof we'll be working with the series $\displaystyle\sum_{n=1}^{\infty} a_n$. The original test statement was for a series that started at a general $n = k$ and while the proof can be done for that it will be easier if we assume that the series starts at $n = 1$.

Another way of dealing with the $n = k$ is we could do an index shift and start the series at $n = 1$ and then do the Integral Test. Either way proving the test for $n = 1$ will be sufficient.

Let's start off and estimate the area under the curve on the interval $[1, n]$ and we'll underestimate the area by taking rectangles of width one and whose height is the right endpoint. This gives the following figure.



Now, note that,

$$f(2) = a_2 \qquad f(3) = a_3 \qquad \cdots \qquad f(n) = a_n$$

The approximate area is then,

$$A \approx (1)\, f(2) + (1)\, f(3) + \cdots + (1)\, f(n) = a_2 + a_3 + \cdots a_n$$

and we know that this underestimates the actual area so,

$$\sum_{i=2}^{n} a_i = a_2 + a_3 + \cdots a_n < \int_1^n f(x)\,dx$$

Now, let's suppose that $\int_1^\infty f(x)\,dx$ is convergent and so $\int_1^\infty f(x)\,dx$ must have a finite value. Also, because $f(x)$ is positive we know that,

$$\int_1^n f(x)\,dx < \int_1^\infty f(x)\,dx$$

This in turn means that,

$$\sum_{i=2}^{n} a_i < \int_1^n f(x)\,dx < \int_1^\infty f(x)\,dx$$

Our series starts at $n = 1$ so this isn't quite what we need. However, that's easy enough to deal with.

$$\sum_{i=1}^{n} a_i = a_1 + \sum_{i=2}^{n} a_i < a_1 + \int_1^\infty f(x)\,dx = M$$

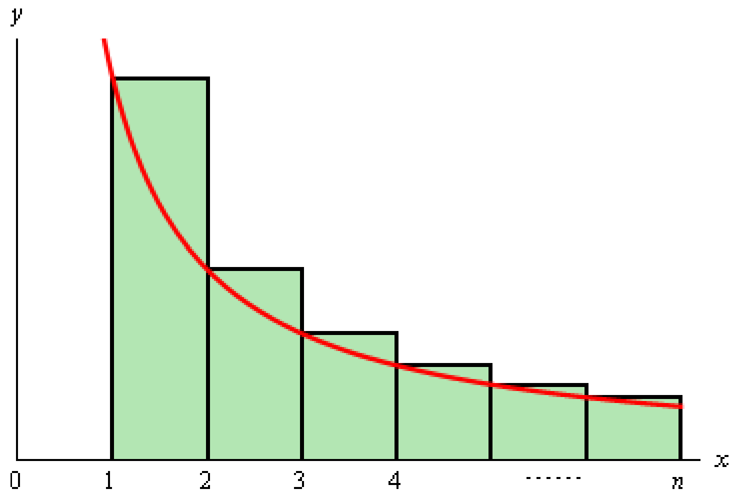
So, just what has this told us? Well we now know that the sequence of partial sums, $s_n = \sum_{i=1}^{n} a_i$ are bounded above by *M*.

Next, because the terms are positive we also know that,

$$s_n \le s_n + a_{n+1} = \sum_{i=1}^{n} a_i + a_{n+1} = \sum_{i=1}^{n+1} a_i = s_{n+1} \qquad \Rightarrow \qquad s_n \le s_{n+1}$$

and so the sequence $\{s_n\}_{n=1}^\infty$ is also an increasing sequence. So, we now know that the sequence of partial sums $\{s_n\}_{n=1}^\infty$ converges and hence our series $\sum_{n=1}^{\infty} a_n$ is convergent.

So, the first part of the test is proven. The second part is somewhat easier. This time let's overestimate the area under the curve by using the left endpoints of interval for the height of the rectangles as shown below.



In this case the area is approximately,

$$A \approx (1)\,f(1) + (1)\,f(2) + \cdots + (1)\,f(n-1) = a_1 + a_2 + \cdots a_{n-1}$$

Since we know this overestimates the area we also then know that,

$$s_{n-1} = \sum_{i=1}^{n-1} a_i = a_1 + a_2 + \cdots a_{n-1} > \int_1^{n-1} f(x)\,dx$$

Now, suppose that $\int_1^{\infty} f(x)\,dx$ is divergent. In this case this means that $\int_1^{n} f(x)\,dx \to \infty$ as $n \to \infty$ because $f(x) \geq 0$. However, because $n-1 \to \infty$ as $n \to \infty$ we also know that $\int_1^{n-1} f(x)\,dx \to \infty$.

Therefore, since $s_{n-1} > \int_1^{n-1} f(x)\,dx$ we know that as $n \to \infty$ we must have $s_{n-1} \to \infty$. This in turn tells us that $s_n \to \infty$ as $n \to \infty$.

So, we now know that the sequence of partial sums, $\left\{s_n\right\}_{n=1}^{\infty}$, is a divergent sequence and so $\sum_{n=1}^{\infty} a_n$ is a divergent series.

■

It is important to note before leaving this section that in order to use the Integral Test the series terms MUST eventually be decreasing and positive. If they are not then the test doesn't work. Also remember that the test only determines the convergence of a series and does NOT give the value of the series.

Comparison Test / Limit Comparison Test

In the previous section we saw how to relate a series to an improper integral to determine the convergence of a series. While the integral test is a nice test, it does force us to do improper integrals which aren't always easy and in some cases may be impossible to determine the convergence of.

For instance consider the following series.

$$\sum_{n=0}^{\infty} \frac{1}{3^n + n}$$

In order to use the Integral Test we would have to integrate

$$\int_0^{\infty} \frac{1}{3^x + x}\, dx$$

and I'm not even sure if it's possible to do this integral. Nicely enough for us there is another test that we can use on this series that will be much easier to use.

First, let's note that the series terms are positive. As with the Integral Test that will be important in this section. Next let's note that we must have $x > 0$ since we are integrating on the interval $0 \le x < \infty$. Likewise, regardless of the value of *x* we will always have $3^x > 0$. So, if we drop the *x* from the denominator the denominator will get smaller and hence the whole fraction will get larger. So,

$$\frac{1}{3^n + n} < \frac{1}{3^n}$$

Now,

$$\sum_{n=0}^{\infty} \frac{1}{3^n}$$

is a geometric series and we know that since $\left| r \right| = \left| \frac{1}{3} \right| < 1$ the series will converge and its value will be,

$$\sum_{n=0}^{\infty} \frac{1}{3^n} = \frac{1}{1 - \frac{1}{3}} = \frac{3}{2}$$

Now, if we go back to our original series and write down the partial sums we get,

$$s_n = \sum_{i=0}^{n} \frac{1}{3^i + i}$$

Since all the terms are positive adding a new term will only make the number larger and so the sequence of partial sums must be an increasing sequence.

$$s_n = \sum_{i=0}^{n} \frac{1}{3^i + i} < \sum_{i=0}^{n+1} \frac{1}{3^i + i} = s_{n+1}$$

Then since,

$$\frac{1}{3^n + n} < \frac{1}{3^n}$$

and because the terms in these two sequences are positive we can also say that,

$$s_n = \sum_{i=0}^{n} \frac{1}{3^i + i} < \sum_{i=0}^{n} \frac{1}{3^i} < \sum_{i=0}^{\infty} \frac{1}{3^n} = \frac{3}{2} \qquad \Rightarrow \qquad s_n < \frac{3}{2}$$

Therefore, the sequence of partial sums is also a bounded sequence. Then from the second section on sequences we know that a monotonic and bounded sequence is also convergent.

So, the sequence of partial sums of our series is a convergent sequence. This means that the series itself,

$$\sum_{n=0}^{\infty} \frac{1}{3^n + n}$$

is also convergent.

So, what did we do here? We found a series whose terms were always larger than the original series terms and this new series was also convergent. Then since the original series terms were positive (very important) this meant that the original series was also convergent.

To show that a series (with only positive terms) was divergent we could go through a similar argument and find a new divergent series whose terms are always smaller than the original series. In this case the original series would have to take a value larger than the new series. However, since the new series is divergent its value will be infinite. This means that the original series must also be infinite and hence divergent.

We can summarize all this in the following test.

Comparison Test

Suppose that we have two series $\sum a_n$ and $\sum b_n$ with $a_n, b_n \geq 0$ for all *n* and $a_n \leq b_n$ for all *n*. Then,

1. If $\sum b_n$ is convergent then so is $\sum a_n$.
2. If $\sum a_n$ is divergent then so is $\sum b_n$.

In other words, we have two series of positive terms and the terms of one of the series is always larger than the terms of the other series. Then if the larger series is convergent the smaller series must also be convergent. Likewise, if the smaller series is divergent then the larger series must also be divergent. Note as well that in order to apply this test we need both series to start at the same place.

A formal proof of this test is at the end of this section.

Do not misuse this test. Just because the smaller of the two series converges does not say anything about the larger series. The larger series may still diverge. Likewise, just because we know that the larger of two series diverges we can't say that the smaller series will also diverge! Be very careful in using this test

Recall that we had a similar test for improper integrals back when we were looking at integration techniques. So, if you could use the comparison test for improper integrals you can use the comparison test for series as they are pretty much the same idea.

Note as well that the requirement that $a_n, b_n \geq 0$ and $a_n \leq b_n$ really only need to be true eventually. In other words, if a couple of the first terms are negative or $a_n \not\leq b_n$ for a couple of the first few terms we're okay. As long as we eventually reach a point where $a_n, b_n \geq 0$ and $a_n \leq b_n$ for all sufficiently large *n* the test will work.

To see why this is true let's suppose that the series start at $n = k$ and that the conditions of the test are only true for for $n \geq N+1$ and for $k \leq n \leq N$ at least one of the conditions is not true. If we then look at $\sum a_n$ (the same thing could be done for $\sum b_n$) we get,

$$\sum_{n=k}^{\infty} a_n = \sum_{n=k}^{N} a_n + \sum_{n=N+1}^{\infty} a_n$$

The first series is nothing more than a finite sum (no matter how large *N* is) of finite terms and so will be finite. So the original series will be convergent/divergent only if the second infinite series on the right is convergent/divergent and the test can be done on the second series as it satisfies the conditions of the test.

Let's take a look at some examples.

Example 1 Determine if the following series is convergent or divergent.

$$\sum_{n=1}^{\infty} \frac{n}{n^2 - \cos^2(n)}$$

Solution
Since the cosine term in the denominator doesn't get too large we can assume that the series terms will behave like,

$$\frac{n}{n^2} = \frac{1}{n}$$

which, as a series, will diverge. So, from this we can guess that the series will probably diverge and so we'll need to find a smaller series that will also diverge.

Recall that from the comparison test with improper integrals that we determined that we can make a fraction smaller by either making the numerator smaller or the denominator larger. In this case the two terms in the denominator are both positive. So, if we drop the cosine term we will in fact be making the denominator larger since we will no longer be subtracting off a positive quantity. Therefore,

$$\frac{n}{n^2 - \cos^2(n)} > \frac{n}{n^2} = \frac{1}{n}$$

Then, since

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges (it's harmonic or the *p*-series test) by the Comparison Test our original series must also diverge.

Example 2 Determine if the following series converges or diverges.

$$\sum_{n=1}^{\infty} \frac{n^2 + 2}{n^4 + 5}$$

Solution

In this case the "+2" and the "+5" don't really add anything to the series and so the series terms should behave pretty much like

$$\frac{n^2}{n^4} = \frac{1}{n^2}$$

which will converge as a series. Therefore, we can guess that the original series will converge and we will need to find a larger series which also converges.

This means that we'll either have to make the numerator larger or the denominator smaller. We can make the denominator smaller by dropping the "+5". Doing this gives,

$$\frac{n^2 + 2}{n^4 + 5} < \frac{n^2 + 2}{n^4}$$

At this point, notice that we can't drop the "+2" from the numerator since this would make the term smaller and that's not what we want. However, this is actually the furthest that we need to go. Let's take a look at the following series.

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{n^2 + 2}{n^4} &= \sum_{n=1}^{\infty} \frac{n^2}{n^4} + \sum_{n=1}^{\infty} \frac{2}{n^4} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^2} + \sum_{n=1}^{\infty} \frac{2}{n^4}\end{aligned}$$

As shown, we can write the series as a sum of two series and both of these series are convergent by the *p*-series test. Therefore, since each of these series are convergent we know that the sum,

$$\sum_{n=1}^{\infty} \frac{n^2 + 2}{n^4}$$

is also a convergent series. Recall that the sum of two convergent series will also be convergent.

Now, since the terms of this series are larger than the terms of the original series we know that the original series must also be convergent by the Comparison Test.

The comparison test is a nice test that allows us to do problems that either we couldn't have done with the integral test or at the best would have been very difficult to do with the integral test. That doesn't mean that it doesn't have problems of its own.

Consider the following series.

$$\sum_{n=0}^{\infty} \frac{1}{3^n - n}$$

This is not much different from the first series that we looked at. The original series converged because the 3^n gets very large very fast and will be significantly larger than the n. Therefore, the n doesn't really affect the convergence of the series in that case. The fact that we are now subtracting the n off instead of adding the n on really shouldn't change the convergence. We can say this because the 3^n gets very large very fast and the fact that we're subtracting n off won't really change the size of this term for all sufficiently large values of n.

So, we would expect this series to converge. However, the comparison test won't work with this series. To use the comparison test on this series we would need to find a larger series that we could easily determine the convergence of. In this case we can't do what we did with the original series. If we drop the n we will make the denominator larger (since the n was subtracted off) and so the fraction will get smaller and just like when we looked at the comparison test for improper integrals knowing that the smaller of two series converges does not mean that the larger of the two will also converge.

So, we will need something else to do help us determine the convergence of this series. The following variant of the comparison test will allow us to determine the convergence of this series.

Limit Comparison Test

Suppose that we have two series $\sum a_n$ and $\sum b_n$ with $a_n \geq 0, b_n > 0$ for all n. Define,

$$c = \lim_{n\to\infty} \frac{a_n}{b_n}$$

If c is positive (*i.e.* $c > 0$) and is finite (*i.e.* $c < \infty$) then either both series converge or both series diverge.

The proof of this test is at the end of this section.

Note that it doesn't really matter which series term is in the numerator for this test, we could just have easily defined c as,

$$c = \lim_{n\to\infty} \frac{b_n}{a_n}$$

and we would get the same results. To see why this is, consider the following two definitions.

$$c = \lim_{n\to\infty} \frac{a_n}{b_n} \qquad\qquad \overline{c} = \lim_{n\to\infty} \frac{b_n}{a_n}$$

Start with the first definition and rewrite it as follows, then take the limit.

$$c = \lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{1}{\frac{b_n}{a_n}} = \frac{1}{\lim\limits_{n\to\infty} \frac{b_n}{a_n}} = \frac{1}{\overline{c}}$$

In other words, if c is positive and finite then so is $\overline{c}$ and if $\overline{c}$ is positive and finite then so is c. Likewise if $\overline{c} = 0$ then $c = \infty$ and if $\overline{c} = \infty$ then $c = 0$. Both definitions will give the same results from the test so don't worry about which series terms should be in the numerator and which should be in the denominator. Choose this to make the limit easy to compute.

Also, this really is a comparison test in some ways. If c is positive and finite this is saying that both of the series terms will behave in generally the same fashion and so we can expect the series themselves to also behave in a similar fashion. If $c = 0$ or $c = \infty$ we can't say this and so the test fails to give any information.

The limit in this test will often be written as,

$$c = \lim_{n \to \infty} a_n \cdot \frac{1}{b_n}$$

since often both terms will be fractions and this will make the limit easier to deal with.

Let's see how this test works.

Example 3 Determine if the following series converges or diverges.

$$\sum_{n=0}^{\infty} \frac{1}{3^n - n}$$

Solution

To use the limit comparison test we need to find a second series that we can determine the convergence of easily and has what we assume is the same convergence as the given series. On top of that we will need to choose the new series in such a way as to give us an easy limit to compute for c.

We've already guessed that this series converges and since it's vaguely geometric let's use

$$\sum_{n=0}^{\infty} \frac{1}{3^n}$$

as the second series. We know that this series converges and there is a chance that since both series have the 3^n in it the limit won't be too bad.

Here's the limit.

$$\begin{aligned} c &= \lim_{n \to \infty} \frac{1}{3^n} \frac{3^n - n}{1} \\ &= \lim_{n \to \infty} 1 - \frac{n}{3^n} \end{aligned}$$

Now, we'll need to use L'Hospital's Rule on the second term in order to actually evaluate this limit.

$$\begin{aligned} c &= 1 - \lim_{n \to \infty} \frac{1}{3^n \ln(3)} \\ &= 1 \end{aligned}$$

So, c is positive and finite so by the Comparison Test both series must converge since

$$\sum_{n=0}^{\infty} \frac{1}{3^n}$$

converges.

Example 4 Determine if the following series converges or diverges.

$$\sum_{n=2}^{\infty} \frac{4n^2 + n}{\sqrt[3]{n^7 + n^3}}$$

Solution

Fractions involving only polynomials or polynomials under radicals will behave in the same way as the largest power of *n* will behave in the limit. So, the terms in this series should behave as,

$$\frac{n^2}{\sqrt[3]{n^7}} = \frac{n^2}{n^{\frac{7}{3}}} = \frac{1}{n^{\frac{1}{3}}}$$

and as a series this will diverge by the *p*-series test. In fact, this would make a nice choice for our second series in the limit comparison test so let's use it.

$$\lim_{n\to\infty} \frac{4n^2 + n}{\sqrt[3]{n^7 + n^3}} \frac{n^{\frac{1}{3}}}{1} = \lim_{n\to\infty} \frac{4n^{\frac{7}{3}} + n^{\frac{4}{3}}}{\sqrt[3]{n^7\left(1 + \frac{1}{n^4}\right)}}$$

$$= \lim_{n\to\infty} \frac{n^{\frac{7}{3}}\left(4 + \frac{1}{n}\right)}{n^{\frac{7}{3}}\sqrt[3]{1 + \frac{1}{n^4}}}$$

$$= \frac{4}{\sqrt[3]{1}} = 4 = c$$

So, *c* is positive and finite and so both limits will diverge since

$$\sum_{n=2}^{\infty} \frac{1}{n^{\frac{1}{3}}}$$

diverges.

Finally, to see why we need *c* to be positive and finite (*i.e.* $c \neq 0$ and $c \neq \infty$) consider the following two series.

$$\sum_{n=1}^{\infty} \frac{1}{n} \qquad\qquad \sum_{n=1}^{\infty} \frac{1}{n^2}$$

The first diverges and the second converges.

Now compute each of the following limits.

$$\lim_{n\to\infty} \frac{1}{n} \bullet \frac{n^2}{1} = \lim_{n\to\infty} n = \infty \qquad\qquad \lim_{n\to\infty} \frac{1}{n^2} \bullet \frac{n}{1} = \lim_{n\to\infty} \frac{1}{n} = 0$$

In the first case the limit from the limit comparison test yields $c = \infty$ and in the second case the limit yields $c = 0$. Clearly, both series do not have the same convergence.

Note however, that just because we get $c = 0$ or $c = \infty$ doesn't mean that the series will have the opposite convergence. To see this consider the series,

$$\sum_{n=1}^{\infty} \frac{1}{n^3} \qquad \qquad \sum_{n=1}^{\infty} \frac{1}{n^2}$$

Both of these series converge and here are the two possible limits that the limit comparison test uses.

$$\lim_{n \to \infty} \frac{1}{n^3} \bullet \frac{n^2}{1} = \lim_{n \to \infty} \frac{1}{n} = 0 \qquad \qquad \lim_{n \to \infty} \frac{1}{n^2} \bullet \frac{n^3}{1} = \lim_{n \to \infty} n = \infty$$

So, even though both series had the same convergence we got both $c = 0$ and $c = \infty$.

The point of all of this is to remind us that if we get $c = 0$ or $c = \infty$ from the limit comparison test we will know that we have chosen the second series incorrectly and we'll need to find a different choice in order to get any information about the convergence of the series.

We'll close out this section with proofs of the two tests.

Proof of Comparison Test

The test statement did not specify where each series should start. We only need to require that they start at the same place so to help with the proof we'll assume that the series start at $n = 1$. If the series don't start at $n = 1$ the proof can be redone in exactly the same manner or you could use an index shift to start the series at $n = 1$ and then this proof will apply.

We'll start off with the partial sums of each series.

$$s_n = \sum_{i=1}^{n} a_i \qquad \qquad t_n = \sum_{i=1}^{n} b_i$$

Let's notice a couple of nice facts about these two partial sums. First, because $a_n, b_n \ge 0$ we know that,

$$s_n \le s_n + a_{n+1} = \sum_{i=1}^{n} a_i + a_{n+1} = \sum_{i=1}^{n+1} a_i = s_{n+1} \qquad \Rightarrow \qquad s_n \le s_{n+1}$$

$$t_n \le t_n + b_{n+1} = \sum_{i=1}^{n} b_i + b_{n+1} = \sum_{i=1}^{n+1} b_i = t_{n+1} \qquad \Rightarrow \qquad t_n \le t_{n+1}$$

So, both partial sums form increasing sequences.

Also, because $a_n \le b_n$ for all n we know that we must have $s_n \le t_n$ for all n.

With these preliminary facts out of the way we can proceed with the proof of the test itself.

Let's start out by assuming that $\sum_{n=1}^{\infty} b_n$ is a convergent series. Since $b_n \ge 0$ we know that,

$$t_n = \sum_{i=1}^{n} b_i \le \sum_{i=1}^{\infty} b_i$$

However, we also have established that $s_n \le t_n$ for all *n* and so for all *n* we also have,

$$s_n \le \sum_{i=1}^{\infty} b_i$$

Finally since $\sum_{n=1}^{\infty} b_n$ is a convergent series it must have a finite value and so the partial sums, s_n are bounded above. Therefore, from the second section on sequences we know that a monotonic and bounded sequence is also convergent and so $\{s_n\}_{n=1}^{\infty}$ is a convergent sequence and so $\sum_{n=1}^{\infty} a_n$ is convergent.

Next, let's assume that $\sum_{n=1}^{\infty} a_n$ is divergent. Because $a_n \ge 0$ we then know that we must have $s_n \to \infty$ as $n \to \infty$. However, we also know that for all *n* we have $s_n \le t_n$ and therefore we also know that $t_n \to \infty$ as $n \to \infty$.

So, $\{t_n\}_{n=1}^{\infty}$ is a divergent sequence and so $\sum_{n=1}^{\infty} b_n$ is divergent.

■

Proof of Limit Comparison Test

Because $0 < c < \infty$ we can find two positive and finite numbers, *m* and *M*, such that $m < c < M$. Now, because $c = \lim_{n\to\infty} \frac{a_n}{b_n}$ we know that for large enough *n* the quotient $\frac{a_n}{b_n}$ must be close to *c* and so there must be a positive integer *N* such that if $n > N$ we also have,

$$m < \frac{a_n}{b_n} < M$$

Multiplying through by b_n gives,

$$mb_n < a_n < Mb_n$$

provided $n > N$.

Now, if $\sum b_n$ diverges then so does $\sum mb_n$ and so since $mb_n < a_n$ for all sufficiently large *n* by the Comparison Test $\sum a_n$ also diverges.

Likewise, if $\sum b_n$ converges then so does $\sum Mb_n$ and since $a_n < Mb_n$ for all sufficiently large *n* by the Comparison Test $\sum a_n$ also converges.

■

Alternating Series Test

The last two tests that we looked at for series convergence have required that all the terms in the series be positive. Of course there are many series out there that have negative terms in them and so we now need to start looking at tests for these kinds of series.

The test that we are going to look into in this section will be a test for alternating series. An **alternating series** is any series, $\sum a_n$, for which the series terms can be written in one of the following two forms.

$$a_n = (-1)^n b_n \qquad\qquad b_n \geq 0$$
$$a_n = (-1)^{n+1} b_n \qquad\qquad b_n \geq 0$$

There are many other ways to deal with the alternating sign, but they can all be written as one of the two forms above. For instance,

$$(-1)^{n+2} = (-1)^n (-1)^2 = (-1)^n$$
$$(-1)^{n-1} = (-1)^{n+1} (-1)^{-2} = (-1)^{n+1}$$

There are of course many others, but they all follow the same basic pattern of reducing to one of the first two forms given. If you should happen to run into a different form than the first two, don't worry about converting it to one of those forms, just be aware that it can be and so the test from this section can be used.

Alternating Series Test

Suppose that we have a series $\sum a_n$ and either $a_n = (-1)^n b_n$ or $a_n = (-1)^{n+1} b_n$ where $b_n \geq 0$ for all n. Then if,

1. $\lim\limits_{n\to\infty} b_n = 0$ and,
2. $\{b_n\}$ is a decreasing sequence

the series $\sum a_n$ is convergent.

A proof of this test is at the end of the section.

There are a couple of things to note about this test. First, unlike the Integral Test and the Comparison/Limit Comparison Test, this test will only tell us when a series converges and not if a series will diverge.

Secondly, in the second condition all that we need to require is that the series terms, b_n will be eventually decreasing. It is possible for the first few terms of a series to increase and still have the test be valid. All that is required is that eventually we will have $b_n \geq b_{n+1}$ for all n after some point.

To see why this is consider the following series,

$$\sum_{n=1}^{\infty} (-1)^n b_n$$

Let's suppose that for $1 \le n \le N$ $\{b_n\}$ is not decreasing and that for $n \ge N+1$ $\{b_n\}$ is decreasing. The series can then be written as,

$$\sum_{n=1}^{\infty} (-1)^n b_n = \sum_{n=1}^{N} (-1)^n b_n + \sum_{n=N+1}^{\infty} (-1)^n b_n$$

The first series is a finite sum (no matter how large *N* is) of finite terms and so we can compute its value and it will be finite. The convergence of the series will depend solely on the convergence of the second (infinite) series. If the second series has a finite value then the sum of two finite values is also finite and so the original series will converge to a finite value. On the other hand if the second series is divergent either because its value is infinite or it doesn't have a value then adding a finite number onto this will not change that fact and so the original series will be divergent.

The point of all this is that we don't need to require that the series terms be decreasing for all *n*. We only need to require that the series terms will eventually be decreasing since we can always strip out the first few terms that aren't actually decreasing and look only at the terms that are actually decreasing.

Note that, in practice, we don't actually strip out the terms that aren't decreasing. All we do is check that eventually the series terms are decreasing and then apply the test.

Let's work a couple of examples.

Example 1 Determine if the following series is convergent or divergent.

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

Solution

First, identify the b_n for the test.

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} \qquad\qquad b_n = \frac{1}{n}$$

Now, all that we need to do is run through the two conditions in the test.

$$\lim_{n \to \infty} b_n = \lim_{n \to \infty} \frac{1}{n} = 0$$

$$b_n = \frac{1}{n} > \frac{1}{n+1} = b_{n+1}$$

Both conditions are met and so by the Alternating Series Test the series must converge.

The series from the previous example is sometimes called the **Alternating Harmonic Series**. Also, the $(-1)^{n+1}$ could be $(-1)^n$ or any other form of alternating sign and we'd still call it an Alternating Harmonic Series.

In the previous example it was easy to see that the series terms decreased since increasing *n* only increased the denominator for the term and hence made the term smaller. In general however we will need to resort to Calculus I techniques to prove the series terms decrease. We'll see an example of this in a bit.

Example 2 Determine if the following series is convergent or divergent.

$$\sum_{n=1}^{\infty} \frac{(-1)^n n^2}{n^2+5}$$

Solution

First, identify the b_n for the test.

$$\sum_{n=1}^{\infty} \frac{(-1)^n n^2}{n^2+5} = \sum_{n=1}^{\infty} (-1)^n \frac{n^2}{n^2+5} \qquad \Rightarrow \qquad b_n = \frac{n^2}{n^2+5}$$

Let's check the conditions.

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} \frac{n^2}{n^2+5} = 1 \neq 0$$

So, the first condition isn't met and so there is no reason to check the second. Since this condition isn't met we'll need to use another test to check convergence. In these cases where the first condition isn't met it is usually best to use the divergence test.

So, the divergence test requires us to compute the following limit.

$$\lim_{n\to\infty} \frac{(-1)^n n^2}{n^2+5}$$

This limit can be somewhat tricky to evaluate. For a second let's consider the following,

$$\lim_{n\to\infty} \frac{(-1)^n n^2}{n^2+5} = \left(\lim_{n\to\infty} (-1)^n\right)\left(\lim_{n\to\infty} \frac{n^2}{n^2+5}\right)$$

Splitting this limit like this can't be done because this operation requires that both limits exist and while the second one does the first clearly does not. However, it does show us how we can at least convince ourselves that the overall limit does not exist (even if it won't be a direct proof of that fact).

So, let's start with,

$$\lim_{n\to\infty} \frac{(-1)^n n^2}{n^2+5} = \lim_{n\to\infty} \left[(-1)^n \frac{n^2}{n^2+5} \right]$$

Now, the second part of this clearly is going to 1 as $n \to \infty$ while the first part just alternates between 1 and -1. So, as $n \to \infty$ the terms are alternating between positive and negative values that are getting closer and closer to 1 and -1 respectively.

In order for limits to exist we know that the terms need to settle down to a single number and since these clearly don't this limit doesn't exist and so by the Divergence Test this series diverges.

Example 3 Determine if the following series is convergent or divergent.

$$\sum_{n=0}^{\infty} \frac{(-1)^{n-3}\sqrt{n}}{n+4}$$

Solution

Notice that in this case the exponent on the "-1" isn't *n* or *n+1*. That won't change how the test works however so we won't worry about that. In this case we have,

$$b_n = \frac{\sqrt{n}}{n+4}$$

so let's check the conditions.

The first is easy enough to check.

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} \frac{\sqrt{n}}{n+4} = 0$$

The second condition requires some work however. It is not immediately clear that these terms will decrease. Increasing *n* to *n+1* will increase both the numerator and the denominator. Increasing the numerator says the term should also increase while increasing the denominator says that the term should decrease. Since it's not clear which of these will win out we will need to resort to Calculus I techniques to show that the terms decrease.

Let's start with the following function and its derivative.

$$f(x) = \frac{\sqrt{x}}{x+4} \qquad\qquad f'(x) = \frac{4-x}{2\sqrt{x}\,(x+4)^2}$$

Now, there are three critical points for this function, $x=-4$, $x=0$, and $x=4$. The first is outside the bound of our series so we won't need to worry about that one. Using the test points,

$$f'(1) = \frac{3}{50} \qquad\qquad f'(5) = -\frac{\sqrt{5}}{810}$$

and so we can see that the function in increasing on $0 \le x \le 4$ and decreasing on $x \ge 4$. Therefore, since $f(n) = b_n$ we know as well that the b_n are also increasing on $0 \le n \le 4$ and decreasing on $n \ge 4$.

The b_n are then eventually decreasing and so the second condition is met.

Both conditions are met and so by the Alternating Series Test the series must be converging.

As the previous example has shown, we sometimes need to do a fair amount of work to show that the terms are decreasing. Do not just make the assumption that the terms will be decreasing and let it go at that.

Let's do one more example just to make a point.

Example 4 Determine if the following series is convergent or divergent.

$$\sum_{n=2}^{\infty}\frac{\cos(n\pi)}{\sqrt{n}}$$

Solution
The point of this problem is really just to acknowledge that it is in fact an alternating series. To see this we need to acknowledge that,

$$\cos(n\pi) = (-1)^n$$

and so the series is really,

$$\sum_{n=2}^{\infty}\frac{\cos(n\pi)}{\sqrt{n}} = \sum_{n=2}^{\infty}\frac{(-1)^n}{\sqrt{n}} \qquad \Rightarrow \qquad b_n = \frac{1}{\sqrt{n}}$$

Checking the two condition gives,

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty}\frac{1}{\sqrt{n}} = 0$$

$$b_n = \frac{1}{\sqrt{n}} > \frac{1}{\sqrt{n+1}} = b_{n+1}$$

The two conditions of the test are met and so by the Alternating Series Test the series is convergent.

It should be pointed out that the rewrite we did in previous example only works because n is an integer and because of the presence of the π. Without the π we couldn't do this and if n wasn't guaranteed to be an integer we couldn't do this.

Let's close this section out with a proof of the Alternating Series Test.

Proof of Alternating Series Test

Without loss of generality we can assume that the series starts at $n=1$. If not we could modify the proof below to meet the new starting place or we could do an index shift to get the series to start at $n=1$.

First, notice that because the terms of the sequence are decreasing for any two successive terms we can say,

$$b_n - b_{n+1} \geq 0$$

Now, let's take a look at the even partial sums.

$$\begin{aligned}
s_2 &= b_1 - b_2 \geq 0 & & \\
s_4 &= b_1 - b_2 + b_3 - b_4 = s_2 + b_3 - b_4 \geq s_2 & & \text{because } b_3 - b_4 \geq 0 \\
s_6 &= s_4 + b_5 - b_6 \geq s_4 & & \text{because } b_5 - b_6 \geq 0 \\
&\vdots & & \\
s_{2n} &= s_{2n-2} + b_{2n-1} - b_{2n} \geq s_{2n-2} & & \text{because } b_{2n-1} - b_{2n} \geq 0
\end{aligned}$$

So, $\{s_{2n}\}$ is an increasing sequence.

Next, we can also write the general term as,

$$\begin{aligned} s_{2n} &= b_1 - b_2 + b_3 - b_4 + b_5 + \cdots - b_{2n-2} + b_{2n-1} - b_{2n} \\ &= b_1 - (b_2 - b_3) - (b_4 - b_5) + \cdots - (b_{2n-2} - b_{2n-1}) - b_{2n} \end{aligned}$$

Each of the quantities in parenthesis are positive and by assumption we know that b_{2n} is also positive. So, this tells us that $s_{2n} \le b_1$ for all *n*.

We now know that $\{s_{2n}\}$ is an increasing sequence that is bounded above and so we know that it must also converge. So, let's assume that its limit is *s* or,

$$\lim_{n\to\infty} s_{2n} = s$$

Next, we can quickly determine the limit of the sequence of odd partial sums, $\{s_{2n+1}\}$, as follows,

$$\lim_{n\to\infty} s_{2n+1} = \lim_{n\to\infty} (s_{2n} + b_{2n+1}) = \lim_{n\to\infty} s_{2n} + \lim_{n\to\infty} b_{2n+1} = s + 0 = s$$

So, we now know that both $\{s_{2n}\}$ and $\{s_{2n+1}\}$ are convergent sequences and they both have the same limit and so we also know that $\{s_n\}$ is a convergent sequence with a limit of *s*. This in turn tells us that $\sum a_n$ is convergent.

■

Absolute Convergence

When we first talked about series convergence we briefly mentioned a stronger type of convergence but didn't do anything with it because we didn't have any tools at our disposal that we could use to work problems involving it. We now have some of those tools so it's now time to talk about absolute convergence in detail.

First, let's go back over the definition of absolute convergence.

Definition

A series $\sum a_n$ is called **absolutely convergent** if $\sum |a_n|$ is convergent. If $\sum a_n$ is convergent and $\sum |a_n|$ is divergent we call the series **conditionally convergent**.

We also have the following fact about absolute convergence.

Fact

If $\sum a_n$ is absolutely convergent then it is also convergent.

Proof

First notice that $|a_n|$ is either a_n or it is $-a_n$ depending on its sign. This means that we can then say,

$$0 \le a_n + |a_n| \le 2|a_n|$$

Now, since we are assuming that $\sum |a_n|$ is convergent then $\sum 2|a_n|$ is also convergent since we can just factor the 2 out of the series and 2 times a finite value will still be finite. This however allows us to use the Comparison Test to say that $\sum a_n + |a_n|$ is also a convergent series.

Finally, we can write,

$$\sum a_n = \sum a_n + |a_n| - \sum |a_n|$$

and so $\sum a_n$ is the difference of two convergent series and so is also convergent.

■

This fact is one of the ways in which absolute convergence is a "stronger" type of convergence. Series that are absolutely convergent are guaranteed to be convergent. However, series that are convergent may or may not be absolutely convergent.

Let's take a quick look at a couple of examples of absolute convergence.

Example 1 Determine if each of the following series are absolute convergent, conditionally convergent or divergent.

(a) $\displaystyle\sum_{n=1}^{\infty}\frac{\left(-1\right)^n}{n}$ [Solution]

(b) $\displaystyle\sum_{n=1}^{\infty}\frac{\left(-1\right)^{n+2}}{n^2}$ [Solution]

(c) $\displaystyle\sum_{n=1}^{\infty}\frac{\sin n}{n^3}$ [Solution]

Solution

(a) $\displaystyle\sum_{n=1}^{\infty}\frac{\left(-1\right)^n}{n}$

This is the alternating harmonic series and we saw in the last section that it is a convergent series so we don't need to check that here. So, let's see if it is an absolutely convergent series. To do this we'll need to check the convergence of.

$$\sum_{n=1}^{\infty}\left|\frac{\left(-1\right)^n}{n}\right| = \sum_{n=1}^{\infty}\frac{1}{n}$$

This is the harmonic series and we know from the integral test section that it is divergent.

Therefore, this series is not absolutely convergent. It is however conditionally convergent since the series itself does converge.

[Return to Problems]

(b) $\displaystyle\sum_{n=1}^{\infty}\frac{\left(-1\right)^{n+2}}{n^2}$

In this case let's just check absolute convergence first since if it's absolutely convergent we won't need to bother checking convergence as we will get that for free.

$$\sum_{n=1}^{\infty}\left|\frac{\left(-1\right)^{n+2}}{n^2}\right| = \sum_{n=1}^{\infty}\frac{1}{n^2}$$

This series is convergent by the *p*-series test and so the series is absolute convergent. Note that this does say as well that it's a convergent series.

[Return to Problems]

(c) $\displaystyle\sum_{n=1}^{\infty}\frac{\sin n}{n^3}$

In this part we need to be a little careful. First, this is NOT an alternating series and so we can't use any tools from that section.

What we'll do here is check for absolute convergence first again since that will also give convergence. This means that we need to check the convergence of the following series.

$$\sum_{n=1}^{\infty}\left|\frac{\sin n}{n^3}\right| = \sum_{n=1}^{\infty}\frac{|\sin n|}{n^3}$$

To do this we'll need to note that

$$-1 \le \sin n \le 1 \qquad \Rightarrow \qquad |\sin n| \le 1$$

and so we have,

$$\frac{|\sin n|}{n^3} \le \frac{1}{n^3}$$

Now we know that

$$\sum_{n=1}^{\infty}\frac{1}{n^3}$$

converges by the *p*-series test and so by the Comparison Test we also know that

$$\sum_{n=1}^{\infty}\frac{|\sin n|}{n^3}$$

converges.

Therefore the original series is absolutely convergent (and hence convergent).

[Return to Problems]

Let's close this section off by recapping a topic we saw earlier. When we first discussed the convergence of series in detail we noted that we can't think of series as an infinite sum because some series can have different sums if we rearrange their terms. In fact, we gave two rearrangements of an Alternating Harmonic series that gave two different values. We closed that section off with the following fact,

Facts

Given the series $\sum a_n$,

3. If $\sum a_n$ is absolutely convergent and its value is *s* then any rearrangement of $\sum a_n$ will also have a value of *s*.

4. If $\sum a_n$ is conditionally convergent and *r* is any real number then there is a rearrangement of $\sum a_n$ whose value will be *r*.

Now that we've got the tools under our belt to determine absolute and conditional convergence we can make a few more comments about this.

First, as we showed above in Example 1a an Alternating Harmonic is conditionally convergent and so no matter what value we chose there is some rearrangement of terms that will give that value. Note as well that this fact does not tell us what that rearrangement must be only that it does exist.

Next, we showed in Example 1b that,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+2}}{n^2}$$

is absolutely convergent and so no matter how we rearrange the terms of this series we'll always get the same value. In fact, it can be shown that the value of this series is,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+2}}{n^2} = -\frac{\pi^2}{12}$$

Ratio Test

In this section we are going to take a look at a test that we can use to see if a series is absolutely convergent or not. Recall that if a series is absolutely convergent then we will also know that it's convergent and so we will often use it to simply determine the convergence of a series.

Before proceeding with the test let's do a quick reminder of factorials. This test will be particularly useful for series that contain factorials (and we will see some in the applications) so let's make sure we can deal with them before we run into them in an example.

If n is an integer such that $n \geq 0$ then n factorial is defined as,

$$\begin{aligned} &n! = n(n-1)(n-2)\cdots(3)(2)(1) && \text{if } n \geq 1 \\ &0! = 1 && \text{by definition} \end{aligned}$$

Let's compute a couple real quick.

$$\begin{aligned} 1! &= 1 \\ 2! &= 2(1) = 2 \\ 3! &= 3(2)(1) = 6 \\ 4! &= 4(3)(2)(1) = 24 \\ 5! &= 5(4)(3)(2)(1) = 120 \end{aligned}$$

In the last computation above, notice that we could rewrite the factorial in a couple of different ways. For instance,

$$\begin{aligned} 5! &= 5\underbrace{(4)(3)(2)(1)}_{4!} = 5\cdot 4! \\ 5! &= 5(4)\underbrace{(3)(2)(1)}_{3!} = 5(4)\cdot 3! \end{aligned}$$

In general we can always "strip out" terms from a factorial as follows.

$$\begin{aligned} n! &= n(n-1)(n-2)\cdots(n-k)\big(n-(k+1)\big)\cdots(3)(2)(1) \\ &= n(n-1)(n-2)\cdots(n-k)\cdot\big(n-(k+1)\big)! \\ &= n(n-1)(n-2)\cdots(n-k)\cdot(n-k-1)! \end{aligned}$$

We will need to do this on occasion so don't forget about it.

Also, when dealing with factorials we need to be very careful with parenthesis. For instance, $(2n)! \neq 2\ n!$ as we can see if we write each of the following factorials out.

$$\begin{aligned} (2n)! &= (2n)(2n-1)(2n-2)\cdots(3)(2)(1) \\ 2\ n! &= 2\big[(n)(n-1)(n-2)\cdots(3)(2)(1)\big] \end{aligned}$$

Again, we will run across factorials with parenthesis so don't drop them. This is often one of the more common mistakes that students make when they first run across factorials.

Okay, we are now ready for the test.

Ratio Test

Suppose we have the series $\sum a_n$. Define,

$$L = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$$

Then,

1. if $L<1$ the series is absolutely convergent (and hence convergent).
2. if $L>1$ the series is divergent.
3. if $L=1$ the series may be divergent, conditionally convergent, or absolutely convergent.

A proof of this test is at the end of the section.

Notice that in the case of $L=1$ the ratio test is pretty much worthless and we would need to resort to a different test to determine the convergence of the series.

Also, the absolute value bars in the definition of L are absolutely required. If they are not there it will be impossible for us to get the incorrect answer.

Let's take a look at some examples.

Example 1 Determine if the following series is convergent or divergent.

$$\sum_{n=1}^{\infty}\frac{(-10)^n}{4^{2n+1}(n+1)}$$

Solution

With this first example let's be a little careful and make sure that we have everything down correctly. Here are the series terms a_n.

$$a_n = \frac{(-10)^n}{4^{2n+1}(n+1)}$$

Recall that to compute a_{n+1} all that we need to do is substitute $n+1$ for all the n's in a_n.

$$a_{n+1} = \frac{(-10)^{n+1}}{4^{2(n+1)+1}((n+1)+1)} = \frac{(-10)^{n+1}}{4^{2n+3}(n+2)}$$

Now, to define L we will use,

$$L = \lim_{n\to\infty}\left|a_{n+1}\cdot\frac{1}{a_n}\right|$$

since this will be a little easier when dealing with fractions as we've got here. So,

$$L = \lim_{n\to\infty}\left|\frac{(-10)^{n+1}}{4^{2n+3}(n+2)}\;\frac{4^{2n+1}(n+1)}{(-10)^n}\right|$$
$$= \lim_{n\to\infty}\left|\frac{-10(n+1)}{4^2(n+2)}\right|$$
$$= \frac{10}{16}\lim_{n\to\infty}\frac{n+1}{n+2}$$
$$= \frac{10}{16} < 1$$

So, $L < 1$ and so by the Ratio Test the series converges absolutely and hence will converge.

As seen in the previous example there is usually a lot of canceling that will happen in these. Make sure that you do this canceling. If you don't do this kind of canceling it can make the limit fairly difficult.

Example 2 Determine if the following series is convergent or divergent.
$$\sum_{n=0}^{\infty}\frac{n!}{5^n}$$

Solution
Now that we've worked one in detail we won't go into quite the detail with the rest of these. Here is the limit.
$$L = \lim_{n\to\infty}\left|\frac{(n+1)!}{5^{n+1}}\frac{5^n}{n!}\right| = \lim_{n\to\infty}\frac{(n+1)!}{5\;n!}$$

In order to do this limit we will need to eliminate the factorials. We simply can't do the limit with the factorials in it. To eliminate the factorials we will recall from our discussion on factorials above that we can always "strip out" terms from a factorial. If we do that with the numerator (in this case because it's the larger of the two) we get,
$$L = \lim_{n\to\infty}\frac{(n+1)\;n!}{5\;n!}$$
at which point we can cancel the $n!$ for the numerator an denominator to get,
$$L = \lim_{n\to\infty}\frac{(n+1)}{5} = \infty > 1$$

So, by the Ratio Test this series diverges.

Example 3 Determine if the following series is convergent or divergent.
$$\sum_{n=2}^{\infty}\frac{n^2}{(2n-1)!}$$

Solution
In this case be careful in dealing with the factorials.

$$
\begin{aligned}
L &= \lim_{n\to\infty}\left|\frac{(n+1)^2}{(2(n+1)-1)!}\frac{(2n-1)!}{n^2}\right| \\
&= \lim_{n\to\infty}\left|\frac{(n+1)^2}{(2n+1)!}\frac{(2n-1)!}{n^2}\right| \\
&= \lim_{n\to\infty}\frac{(n+1)^2}{(2n+1)(2n)(2n-1)!}\frac{(2n-1)!}{n^2} \\
&= \lim_{n\to\infty}\frac{(n+1)^2}{(2n+1)(2n)(n^2)} \\
&= 0 < 1
\end{aligned}
$$

So, by the Ratio Test this series converges absolutely and so converges.

Example 4 Determine if the following series is convergent or divergent.

$$\sum_{n=1}^{\infty}\frac{9^n}{(-2)^{n+1}n}$$

Solution

Do not mistake this for a geometric series. The *n* in the denominator means that this isn't a geometric series. So, let's compute the limit.

$$
\begin{aligned}
L &= \lim_{n\to\infty}\left|\frac{9^{n+1}}{(-2)^{n+2}(n+1)}\frac{(-2)^{n+1}n}{9^n}\right| \\
&= \lim_{n\to\infty}\left|\frac{9n}{(-2)(n+1)}\right| \\
&= \frac{9}{2}\lim_{n\to\infty}\frac{n}{n+1} \\
&= \frac{9}{2} > 1
\end{aligned}
$$

Therefore, by the Ratio Test this series is divergent.

In the previous example the absolute value bars were required to get the correct answer. If we hadn't used them we would have gotten $L = -\frac{9}{2} < 1$ which would have implied a convergent series!

Now, let's take a look at a couple of examples to see what happens when we get $L = 1$. Recall that the ratio test will not tell us anything about the convergence of these series. In both of these examples we will first verify that we get $L = 1$ and then use other tests to determine the convergence.

Example 5 Determine if the following series is convergent or divergent.

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{n^2+1}$$

Solution
Let's first get L.

$$L=\lim_{n\to\infty}\left|\frac{(-1)^{n+1}}{(n+1)^2+1}\,\frac{n^2+1}{(-1)^n}\right|=\lim_{n\to\infty}\frac{n^2+1}{(n+1)^2+1}=1$$

So, as implied earlier we get $L=1$ which means the ratio test is no good for determining the convergence of this series. We will need to resort to another test for this series. This series is an alternating series and so let's check the two conditions from that test.

$$\lim_{n\to\infty}b_n=\lim_{n\to\infty}\frac{1}{n^2+1}=0$$

$$b_n=\frac{1}{n^2+1}>\frac{1}{(n+1)^2+1}=b_{n+1}$$

The two conditions are met and so by the Alternating Series Test this series is convergent. We'll leave it to you to verify this series is also absolutely convergent.

Example 6 Determine if the following series is convergent or divergent.

$$\sum_{n=0}^{\infty}\frac{n+2}{2n+7}$$

Solution
Here's the limit.

$$L=\lim_{n\to\infty}\left|\frac{n+3}{2(n+1)+7}\,\frac{2n+7}{n+2}\right|=\lim_{n\to\infty}\frac{(n+3)(2n+7)}{(2n+9)(n+2)}=1$$

Again, the ratio test tells us nothing here. We can however, quickly use the divergence test on this. In fact that probably should have been our first choice on this one anyway.

$$\lim_{n\to\infty}\frac{n+2}{2n+7}=\frac{1}{2}\neq 0$$

By the Divergence Test this series is divergent.

So, as we saw in the previous two examples if we get $L=1$ from the ratio test the series can be either convergent or divergent.

There is one more thing that we should note about the ratio test before we move onto the next section. The last series was a polynomial divided by a polynomial and we saw that we got $L=1$ from the ratio test. This will always happen with rational expression involving only polynomials or polynomials under radicals. So, in the future it isn't even worth it to try the ratio test on these kinds of problems since we now know that we will get $L=1$.

Also, in the second to last example we saw an example of an alternating series in which the positive term was a rational expression involving polynomials and again we will always get $L = 1$ in these cases.

Let's close the section out with a proof of the Ratio Test.

Proof of Ratio Test

First note that we can assume without loss of generality that the series will start at $n = 1$ as we've done for all our series test proofs.

Let's start off the proof here by assuming that $L < 1$ and we'll need to show that $\sum a_n$ is absolutely convergent. To do this let's first note that because $L < 1$ there is some number *r* such that $L < r < 1$.

Now, recall that,

$$L = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

and because we also have chosen *r* such that $L < r$ there is some *N* such that if $n \ge N$ we will have,

$$\left| \frac{a_{n+1}}{a_n} \right| < r \qquad \Rightarrow \qquad \left| a_{n+1} \right| < r \left| a_n \right|$$

Next, consider the following,

$$\begin{aligned} \left| a_{N+1} \right| &< r \left| a_N \right| \\ \left| a_{N+2} \right| &< r \left| a_{N+1} \right| < r^2 \left| a_N \right| \\ \left| a_{N+3} \right| &< r \left| a_{N+2} \right| < r^3 \left| a_N \right| \\ &\vdots \\ \left| a_{N+k} \right| &< r \left| a_{N+k-1} \right| < r^k \left| a_N \right| \end{aligned}$$

So, for $k = 1, 2, 3, \ldots$ we have $\left| a_{N+k} \right| < r^k \left| a_N \right|$. Just why is this important? Well we can now look at the following series.

$$\sum_{k=0}^{\infty} \left| a_N \right| r^k$$

This is a geometric series and because $0 < r < 1$ we in fact know that it is a convergent series. Also because $\left| a_{N+k} \right| < r^k \left| a_N \right|$ by the Comparison test the series

$$\sum_{n=N+1}^{\infty} \left| a_n \right| = \sum_{k=1}^{\infty} \left| a_{N+k} \right|$$

is convergent. However since,

$$\sum_{n=1}^{\infty} \left| a_n \right| = \sum_{n=1}^{N} \left| a_n \right| + \sum_{n=N+1}^{\infty} \left| a_n \right|$$

we know that $\sum_{n=1}^{\infty} |a_n|$ is also convergent since the first term on the right is a finite sum of finite terms and hence finite. Therefore $\sum_{n=1}^{\infty} a_n$ is absolutely convergent.

Next, we need to assume that $L > 1$ and we'll need to show that $\sum a_n$ is divergent. Recalling that,

$$L = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

and because $L > 1$ we know that there must be some N such that if $n \ge N$ we will have,

$$\left| \frac{a_{n+1}}{a_n} \right| > 1 \qquad \Rightarrow \qquad |a_{n+1}| > |a_n|$$

However, if $|a_{n+1}| > |a_n|$ for all $n \ge N$ then we know that,

$$\lim_{n\to\infty} |a_n| \ne 0$$

because the terms are getting larger and guaranteed to not be negative. This in turn means that,

$$\lim_{n\to\infty} a_n \ne 0$$

Therefore, by the Divergence Test $\sum a_n$ is divergent.

Finally, we need to assume that $L = 1$ and show that we could get a series that has any of the three possibilities. To do this we just need a series for each case. We'll leave the details of checking to you but all three of the following series have $L = 1$ and each one exhibits one of the possibilities.

$\sum_{n=1}^{\infty} \frac{1}{n^2}$	absolutely convergent
$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$	conditionally convergent
$\sum_{n=1}^{\infty} \frac{1}{n}$	divergent

■

Root Test

This is the last test for series convergence that we're going to be looking at. As with the Ratio Test this test will also tell whether a series is absolutely convergent or not rather than simple convergence.

Root Test

Suppose that we have the series $\sum a_n$. Define,

$$L = \lim_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} |a_n|^{\frac{1}{n}}$$

Then,

4. if $L < 1$ the series is absolutely convergent (and hence convergent).
5. if $L > 1$ the series is divergent.
6. if $L = 1$ the series may be divergent, conditionally convergent, or absolutely convergent.

A proof of this test is at the end of the section.

As with the ratio test, if we get $L = 1$ the root test will tell us nothing and we'll need to use another test to determine the convergence of the series. Also note that if $L = 1$ in the Ratio Test then the Root Test will also give $L = 1$.

We will also need the following fact in some of these problems.

Fact

$$\lim_{n\to\infty} n^{\frac{1}{n}} = 1$$

Let's take a look at a couple of examples.

Example 1 Determine if the following series is convergent or divergent.

$$\sum_{n=1}^{\infty} \frac{n^n}{3^{1+2n}}$$

Solution

There really isn't much to these problems other than computing the limit and then using the root test. Here is the limit for this problem.

$$L = \lim_{n\to\infty} \left|\frac{n^n}{3^{1+2n}}\right|^{\frac{1}{n}} = \lim_{n\to\infty} \frac{n}{3^{\frac{1}{n}+2}} = \frac{\infty}{3^2} = \infty > 1$$

So, by the Root Test this series is divergent.

Example 2 Determine if the following series is convergent or divergent.

$$\sum_{n=0}^{\infty} \left(\frac{5n-3n^3}{7n^3+2}\right)^n$$

Solution

Again, there isn't too much to this series.

$$L = \lim_{n\to\infty} \left| \left(\frac{5n - 3n^3}{7n^3 + 2} \right)^n \right|^{\frac{1}{n}} = \lim_{n\to\infty} \left| \frac{5n - 3n^3}{7n^3 + 2} \right| = \left| \frac{-3}{7} \right| = \frac{3}{7} < 1$$

Therefore, by the Root Test this series converges absolutely and hence converges.

Note that we had to keep the absolute value bars on the fraction until we'd taken the limit to get the sign correct.

Example 3 Determine if the following series is convergent or divergent.

$$\sum_{n=3}^{\infty} \frac{(-12)^n}{n}$$

Solution

Here's the limit for this series.

$$L = \lim_{n\to\infty} \left| \frac{(-12)^n}{n} \right|^{\frac{1}{n}} = \lim_{n\to\infty} \frac{12}{n^{\frac{1}{n}}} = \frac{12}{1} = 12 > 1$$

After using the fact from above we can see that the Root Test tells us that this series is divergent.

Proof of Root Test

First note that we can assume without loss of generality that the series will start at $n = 1$ as we've done for all our series test proofs. Also note that this proof is very similiar to the proof of the Ratio Test.

Let's start off the proof here by assuming that $L < 1$ and we'll need to show that $\sum a_n$ is absolutely convergent. To do this let's first note that because $L < 1$ there is some number *r* such that $L < r < 1$.

Now, recall that,

$$L = \lim_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} |a_n|^{\frac{1}{n}}$$

and because we also have chosen *r* such that $L < r$ there is some *N* such that if $n \geq N$ we will have,

$$|a_n|^{\frac{1}{n}} < r \qquad \Rightarrow \qquad |a_n| < r^n$$

Now the series

$$\sum_{n=0}^{\infty} r^n$$

is a geometric series and because $0 < r < 1$ we in fact know that it is a convergent series. Also because $|a_n| < r^n \quad n \geq N$ by the Comparison test the series

$$\sum_{n=N}^{\infty}\left|a_n\right|$$

is convergent. However since,

$$\sum_{n=1}^{\infty}\left|a_n\right| = \sum_{n=1}^{N-1}\left|a_n\right| + \sum_{n=N}^{\infty}\left|a_n\right|$$

we know that $\sum_{n=1}^{\infty}\left|a_n\right|$ is also convergent since the first term on the right is a finite sum of finite terms and hence finite. Therefore $\sum_{n=1}^{\infty} a_n$ is absolutely convergent.

Next, we need to assume that $L > 1$ and we'll need to show that $\sum a_n$ is divergent. Recalling that,

$$L = \lim_{n\to\infty}\sqrt[n]{\left|a_n\right|} = \lim_{n\to\infty}\left|a_n\right|^{\frac{1}{n}}$$

and because $L > 1$ we know that there must be some N such that if $n \ge N$ we will have,

$$\left|a_n\right|^{\frac{1}{n}} > 1 \qquad \Rightarrow \qquad \left|a_n\right| > 1^n = 1$$

However, if $\left|a_n\right| > 1$ for all $n \ge N$ then we know that,

$$\lim_{n\to\infty}\left|a_n\right| \ne 0$$

This in turn means that,

$$\lim_{n\to\infty} a_n \ne 0$$

Therefore, by the Divergence Test $\sum a_n$ is divergent.

Finally, we need to assume that $L = 1$ and show that we could get a series that has any of the three possibilities. To do this we just need a series for each case. We'll leave the details of checking to you but all three of the following series have $L = 1$ and each one exhibits one of the possibilities.

$$\sum_{n=1}^{\infty}\frac{1}{n^2} \qquad \text{absolutely convergent}$$

$$\sum_{n=1}^{\infty}\frac{\left(-1\right)^n}{n} \qquad \text{conditionally convergent}$$

$$\sum_{n=1}^{\infty}\frac{1}{n} \qquad \text{divergent}$$

■

Strategy for Series

Now that we've got all of our tests out of the way it's time to think about organizing all of them into a general set of guidelines to help us determine the convergence of a series.

Note that these are a general set of guidelines and because some series can have more than one test applied to them we will get a different result depending on the path that we take through this set of guidelines. In fact, because more than one test may apply, you should always go completely through the guidelines and identify all possible tests that can be used on a given series. Once this has been done you can identify the test that you feel will be the easiest for you to use.

With that said here is the set of guidelines for determining the convergence of a series.

1. With a quick glance does it look like the series terms don't converge to zero in the limit, *i.e.* does $\lim\limits_{n\to\infty} a_n \neq 0$? If so, use the Divergence Test. Note that you should only do the Divergence Test if a quick glance suggests that the series terms may not converge to zero in the limit.
2. Is the series a *p*-series ($\sum \frac{1}{n^p}$) or a geometric series ($\sum\limits_{n=0}^{\infty} ar^n$ or $\sum\limits_{n=1}^{\infty} ar^{n-1}$)? If so use the fact that *p*-series will only converge if $p > 1$ and a geometric series will only converge if $\left|r\right| < 1$. Remember as well that often some algebraic manipulation is required to get a geometric series into the correct form.
3. Is the series similar to a *p*-series or a geometric series? If so, try the Comparison Test.
4. Is the series a rational expression involving only polynomials or polynomials under radicals (*i.e.* a fraction involving only polynomials or polynomials under radicals)? If so, try the Comparison Test and/or the Limit Comparison Test. Remember however, that in order to use the Comparison Test and the Limit Comparison Test the series terms all need to be positive.
5. Does the series contain factorials or constants raised to powers involving *n*? If so, then the Ratio Test may work. Note that if the series term contains a factorial then the only test that we've got that will work is the Ratio Test.
6. Can the series terms be written in the form $a_n = \left(-1\right)^n b_n$ or $a_n = \left(-1\right)^{n+1} b_n$? If so, then the Alternating Series Test may work.
7. Can the series terms be written in the form $a_n = \left(b_n\right)^n$? If so, then the Root Test may work.
8. If $a_n = f\left(n\right)$ for some positive, decreasing function and $\int_a^{\infty} f\left(x\right) dx$ is easy to evaluate then the Integral Test may work.

Again, remember that these are only a set of guidelines and not a set of hard and fast rules to use when trying to determine the best test to use on a series. If more than one test can be used try to use the test that will be the easiest for you to use and remember that what is easy for someone else may not be easy for you!

Also just so we can put all the tests into one place here is a quick listing of all the tests that we've got.

Divergence Test

If $\lim\limits_{n\to\infty} a_n \neq 0$ then $\sum a_n$ will diverge

Integral Test

Suppose that $f(x)$ is a positive, decreasing function on the interval $[k,\infty)$ and that $f(n) = a_n$ then,

1. If $\int_k^\infty f(x)\,dx$ is convergent so is $\sum\limits_{n=k}^{\infty} a_n$.
2. If $\int_k^\infty f(x)\,dx$ is divergent so is $\sum\limits_{n=k}^{\infty} a_n$.

Comparison Test

Suppose that we have two series $\sum a_n$ and $\sum b_n$ with $a_n, b_n \geq 0$ for all *n* and $a_n \leq b_n$ for all *n*. Then,

1. If $\sum b_n$ is convergent then so is $\sum a_n$.
2. If $\sum a_n$ is divergent then so is $\sum b_n$.

Limit Comparison Test

Suppose that we have two series $\sum a_n$ and $\sum b_n$ with $a_n, b_n \geq 0$ for all *n*. Define,

$$c = \lim_{n\to\infty} \frac{a_n}{b_n}$$

If *c* is positive (*i.e.* $c > 0$) and is finite (*i.e.* $c < \infty$) then either both series converge or both series diverge.

Alternating Series Test

Suppose that we have a series $\sum a_n$ and either $a_n = (-1)^n b_n$ or $a_n = (-1)^{n+1} b_n$ where $b_n \geq 0$ for all *n*. Then if,

1. $\lim\limits_{n\to\infty} b_n = 0$ and,
2. $\{b_n\}$ is eventually a decreasing sequence

the series $\sum a_n$ is convergent

Ratio Test

Suppose we have the series $\sum a_n$. Define,

$$L = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$$

Then,

1. if $L < 1$ the series is absolutely convergent (and hence convergent).
2. if $L > 1$ the series is divergent.
3. if $L = 1$ the series may be divergent, conditionally convergent, or absolutely convergent.

Root Test

Suppose that we have the series $\sum a_n$. Define,

$$L = \lim_{n\to\infty}\sqrt[n]{|a_n|} = \lim_{n\to\infty}|a_n|^{\frac{1}{n}}$$

Then,

1. if $L < 1$ the series is absolutely convergent (and hence convergent).
2. if $L > 1$ the series is divergent.
3. if $L = 1$ the series may be divergent, conditionally convergent, or absolutely convergent.

Estimating the Value of a Series

We have now spent quite a few sections determining the convergence of a series, however, with the exception of geometric and telescoping series, we have not talked about finding the value of a series. This is usually a very difficult thing to do and we still aren't going to talk about how to find the value of a series. What we will do is talk about how to estimate the value of a series. Often that is all that you need to know.

Before we get into how to estimate the value of a series let's remind ourselves how series convergence works. It doesn't make any sense to talk about the value of a series that doesn't converge and so we will be assuming that the series we're working with converges. Also, as we'll see the main method of estimating the value of series will come out of this discussion.

So, let's start with the series $\sum_{n=1}^{\infty} a_n$ (the starting point is not important, but we need a starting point to do the work) and let's suppose that the series converges to s. Recall that this means that if we get the partial sums,

$$s_n = \sum_{i=1}^{n} a_i$$

then they will form a convergent sequence and its limit is s. In other words,

$$\lim_{n\to\infty} s_n = s$$

Now, just what does this mean for us? Well, since this limit converges it means that we can make the partial sums, s_n, as close to s as we want simply by taking n large enough. In other words, if we take n large enough then we can say that,

$$s_n \approx s$$

This is one method of estimating the value of a series. We can just take a partial sum and use that as an estimation of the value of the series. There are now two questions that we should ask about this.

First, how good is the estimation? If we don't have an idea of how good the estimation is then it really doesn't do all that much for us as an estimation.

Secondly, is there any way to make the estimate better? Sometimes we can use this as a starting point and make the estimation better. We won't always be able to do this, but if we can that will be nice.

So, let's start with a general discussion about the determining how good the estimation is. Let's first start with the full series and strip out the first n terms.

$$\sum_{i=1}^{\infty} a_i = \sum_{i=1}^{n} a_i + \sum_{i=n+1}^{\infty} a_i \qquad (1)$$

Note that we converted over to an index of i in order to make the notation consistent with prior notation. Recall that we can use any letter for the index and it won't change the value.

Now, notice that the first series (the n terms that we've stripped out) is nothing more than the partial sum s_n. The second series on the right (the one starting at $i = n+1$) is called the remainder and denoted by R_n. Finally let's acknowledge that we also know the value of the series since we are assuming it's convergent. Taking this notation into account we can rewrite (1) as,

$$s = s_n + R_n$$

We can solve this for the remainder to get,

$$R_n = s - s_n$$

So, the remainder tells us the difference, or error, between the exact value of the series and the value of the partial sum that we are using as the estimation of the value of the series.

Of course we can't get our hands on the actual value of the remainder because we don't have the actual value of the series. However, we can use some of the tests that we've got for convergence to get a pretty good estimate of the remainder provided we make some assumptions about the series. Once we've got an estimate on the value of the remainder we'll also have an idea on just how good a job the partial sum does of estimating the actual value of the series.

There are several tests that will allow us to get estimates of the remainder. We'll go through each one separately.

Integral Test

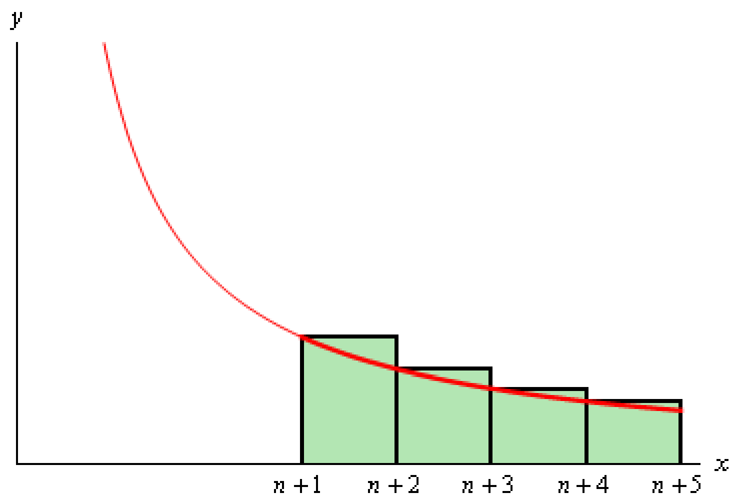
Recall that in this case we will need to assume that the series terms are all positive and will eventually be decreasing. We derived the integral test by using the fact that the series could be thought of as an estimation of the area under the curve of $f(x)$ where $f(n) = a_n$. We can do something similar with the remainder.

As we'll soon see if we can get an upper and lower bound on the value of the remainder we can use these bounds to help us get upper and lower bounds on the value of the series. We can in turn use the upper and lower bounds on the series value to actually estimate the value of the series.

So, let's first recall that the remainder is,

$$R_n = \sum_{i=n+1}^{\infty} a_i = a_{n+1} + a_{n+2} + a_{n+3} + a_{n+4} + \cdots$$

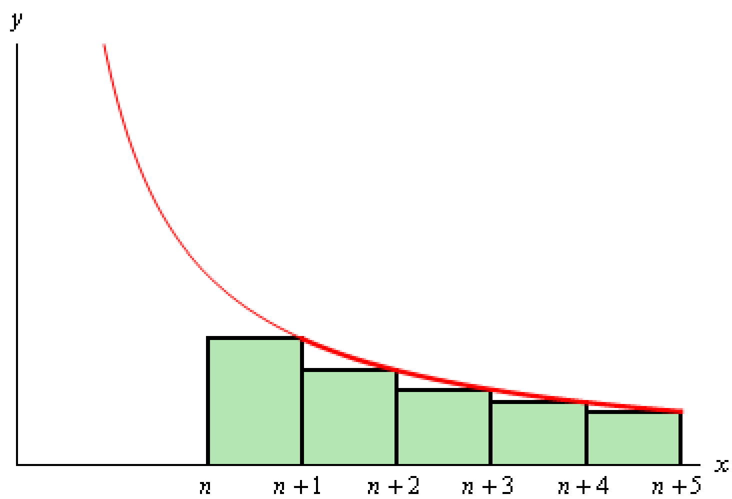
Now, if we start at $x = n+1$, take rectangles of width 1 and use the left endpoint as the height of the rectangle we can estimate the area under $f(x)$ on the interval $[n+1, \infty)$ as shown in the sketch below.



We can see that the remainder, R_n, is the area estimation and it will overestimate the exact area. So, we have the following inequality.

$$R_n \geq \int_{n+1}^{\infty} f(x)\,dx \qquad (2)$$

Next, we could also estimate the area by starting at $x = n$, taking rectangles of width 1 again and then using the right endpoint as the height of the rectangle. This will give an estimation of the area under $f(x)$ on the interval $[n, \infty)$. This is shown in the following sketch.



Again, we can see that the remainder, R_n, is again this estimation and in this case it will underestimate the area. This leads to the following inequality,

$$R_n \leq \int_{n}^{\infty} f(x)\,dx \qquad (3)$$

Combining (2) and (3) gives,

$$\int_{n+1}^{\infty} f(x)\,dx \le R_n \le \int_{n}^{\infty} f(x)\,dx$$

So, provided we can do these integrals we can get both an upper and lower bound on the remainder. This will in turn give us an upper bound and a lower bound on just how good the partial sum, s_n, is as an estimation of the actual value of the series.

In this case we can also use these results to get a better estimate for the actual value of the series as well.

First, we'll start with the fact that

$$s = s_n + R_n$$

Now, if we use (2) we get,

$$s = s_n + R_n \ge s_n + \int_{n+1}^{\infty} f(x)\,dx$$

Likewise if we use (3) we get,

$$s = s_n + R_n \le s_n + \int_{n}^{\infty} f(x)\,dx$$

Putting these two together gives us,

$$s_n + \int_{n+1}^{\infty} f(x)\,dx \le s \le s_n + \int_{n}^{\infty} f(x)\,dx \qquad (4)$$

This gives an upper and a lower bound on the actual value of the series. We could then use as an estimate of the actual value of the series the average of the upper and lower bound.

Let's work an example with this.

Example 1 Using $n = 15$ to estimate the value of $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Solution

First, for comparison purposes, we'll note that the actual value of this series is known to be,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} = 1.644934068$$

Using $n = 15$ let's first get the partial sum.

$$s_{15} = \sum_{i=1}^{15} \frac{1}{i^2} = 1.580440283$$

Note that this is "close" to the actual value in some sense, but isn't really all that close either.

Now, let's compute the integrals. These are fairly simple integrals so we'll leave it to you to verify the values.

$$\int_{15}^{\infty} \frac{1}{x^2}\,dx = \frac{1}{15} \qquad\qquad \int_{16}^{\infty} \frac{1}{x^2}\,dx = \frac{1}{16}$$

Plugging these into (4) gives us,

$$1.580440283 + \frac{1}{16} \le s \le 1.580440283 + \frac{1}{15}$$
$$1.642940283 \le s \le 1.647106950$$

Both the upper and lower bound are now very close to the actual value and if we take the average of the two we get the following estimate of the actual value.

$$s \approx 1.6450236165$$

That is pretty darn close to the actual value.

So, that is how we can use the Integral Test to estimate the value of a series. Let's move on to the next test.

Comparison Test

In this case, unlike with the integral test, we may or may not be able to get an idea of how good a particular partial sum will be as an estimate of the exact value of the series. Much of this will depend on how the comparison test is used.

First, let's remind ourselves on how the comparison test actually works. Given a series $\sum a_n$ let's assume that we've used the comparison test to show that it's convergent. Therefore, we found a second series $\sum b_n$ that converged and $a_n \le b_n$ for all n.

What we want to do is determine how good of a job the partial sum,

$$s_n = \sum_{i=1}^{n} a_i$$

will do in estimating the actual value of the series $\sum a_n$. Again, we will use the remainder to do this. Let's actually write down the remainder for both series.

$$R_n = \sum_{i=n+1}^{\infty} a_i \qquad\qquad T_n = \sum_{i=n+1}^{\infty} b_i$$

Now, since $a_n \le b_n$ we also know that

$$R_n \le T_n$$

When using the comparison test it is often the case that the b_n are fairly nice terms and that we might actually be able to get an idea on the size of T_n. For instance, if our second series is a p-series we can use the results from above to get an upper bound on T_n as follows,

$$R_n \le T_n \le \int_{n}^{\infty} g(x)\,dx \qquad\qquad \text{where } g(n) = b_n$$

Also, if the second series is a geometric series then we will be able to compute T_n exactly.

If we are unable to get an idea of the size of T_n then using the comparison test to help with estimates won't do us much good.

Let's take a look at an example.

Example 2 Using $n = 15$ to estimate the value of $\sum_{n=0}^{\infty} \frac{2^n}{4^n + 1}$.

Solution

To do this we'll first need to go through the comparison test so we can get the second series. So,

$$\frac{2^n}{4^n + 1} \leq \frac{2^n}{4^n} = \left(\frac{1}{2}\right)^n$$

and

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n$$

is a geometric series and converges because $|r| = \frac{1}{2} < 1$.

Now that we've gotten our second series let's get the estimate.

$$s_{15} = \sum_{n=0}^{15} \frac{2^n}{4^n + 1} = 1.383062486$$

So, how good is it? Well we know that,

$$R_{15} \leq T_{15} = \sum_{n=16}^{\infty} \left(\frac{1}{2}\right)^n$$

will be an upper bound for the error between the actual value and the estimate. Since our second series is a geometric series we can compute this directly as follows.

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = \sum_{n=0}^{15} \left(\frac{1}{2}\right)^n + \sum_{n=16}^{\infty} \left(\frac{1}{2}\right)^n$$

The series on the left is in the standard form and so we can compute that directly. The first series on the right has a finite number of terms and so can be computed exactly and the second series on the right is the one that we'd like to have the value for. Doing the work gives,

$$\begin{aligned} \sum_{n=16}^{\infty} \left(\frac{1}{2}\right)^n &= \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n - \sum_{n=0}^{15} \left(\frac{1}{2}\right)^n \\ &= \frac{1}{1 - \left(\frac{1}{2}\right)} - 1.999969482 \\ &= 0.000030518 \end{aligned}$$

So, according to this if we use

$$s \approx 1.383062486$$

as an estimate of the actual value we will be off from the exact value by no more than 0.000030518 and that's not too bad.

In this case it can be shown that

$$\sum_{n=0}^{\infty} \frac{2^n}{4^n+1} = 1.383093004$$

and so we can see that the actual error in our estimation is,

$$\text{Error} = \text{Actual} - \text{Estimate} = 1.383093004 - 1.383062486 = 0.000030518$$

Note that in this case the estimate of the error is actually fairly close (and in fact exactly the same) as the actual error. This will not always happen and so we shouldn't expect that to happen in all cases. The error estimate above is simply the upper bound on the error and the actual error will often be less than this value.

Before moving on to the final part of this section let's again note that we will only be able to determine how good the estimate is using the comparison test if we can easily get our hands on the remainder of the second term. The reality is that we won't always be able to do this.

Alternating Series Test
Both of the methods that we've looked at so far have required the series to contain only positive terms. If we allow series to have negative terms in it the process is usually more difficult. However, with that said there is one case where it isn't too bad. That is the case of an alternating series.

Once again we will start off with a convergent series $\sum a_n = \sum (-1)^n b_n$ which in this case happens to be an alternating series that satisfies the conditions of the alternating series test, so we know that $b_n \geq 0$ for all *n*. Also note that we could have any power on the "-1" we just used *n* for the sake of convenience. We want to know how good of an estimation of the actual series value will the partial sum, s_n, be. As with the prior cases we know that the remainder, R_n, will be the error in the estimation and so if we can get a handle on that we'll know approximately how good the estimation is.

From the proof of the Alternating Series Test we can see that *s* will lie between s_n and s_{n+1} for any *n* and so,

$$\left|s - s_n\right| \leq \left|s_{n+1} - s_n\right| = b_{n+1}$$

Therefore,

$$\left|R_n\right| = \left|s - s_n\right| \leq b_{n+1}$$

We needed absolute value bars because we won't know ahead of time if the estimation is larger or smaller than the actual value and we know that the b_n's are positive.

Let's take a look at an example.

Example 3 Using $n = 15$ to estimate the value of $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$.

Solution

This is an alternating series and it does converge. In this case the exact value is known and so for comparison purposes,

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = -\frac{\pi^2}{12} = -0.8224670336$$

Now, the estimation is,

$$s_{15} = \sum_{n=1}^{15} \frac{(-1)^n}{n^2} = -0.8245417574$$

From the fact above we know that

$$|R_{15}| = |s - s_{15}| \le b_{16} = \frac{1}{16^2} = 0.00390625$$

So, our estimation will have an error of no more than 0.00390625. In this case the exact value is known and so the actual error is,

$$|R_{15}| = |s - s_{15}| = 0.0020747238$$

In the previous example the estimation had only half the estimated error. It will often be the case that the actual error will be less than the estimated error. Remember that this is only an upper bound for the actual error.

Ratio Test

This will be the final case that we're going to look at for estimating series values and we are going to have to put a couple of fairly stringent restrictions on the series terms in order to do the work. One of the main restrictions we're going to make is to assume that the series terms are positive. We'll also be adding on another restriction in a bit.

In this case we've used the ratio test to show that $\sum a_n$ is convergent. To do this we computed

$$L = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

and found that $L < 1$.

As with the previous cases we are going to use the remainder, R_n, to determine how good of an estimation of the actual value the partial sum, s_n, is.

To get an estimate of the remainder let's first define the following sequence,

$$r_n = \frac{a_{n+1}}{a_n}$$

We now have two possible cases.

1. If $\{r_n\}$ is a decreasing sequence and $r_{n+1} < 1$ then,

$$R_n \le \frac{a_{n+1}}{1-r_{n+1}}$$

2. If $\{r_n\}$ is a increasing sequence then,

$$R_n \le \frac{a_{n+1}}{1-L}$$

Proof

Both parts will need the following work so we'll do it first. We'll start with the remainder.

$$\begin{aligned} R_n &= \sum_{i=n+1}^{\infty} a_i = a_{n+1} + a_{n+2} + a_{n+3} + a_{n+4} + \cdots \\ &= a_{n+1}\left(1 + \frac{a_{n+2}}{a_{n+1}} + \frac{a_{n+3}}{a_{n+1}} + \frac{a_{n+4}}{a_{n+1}} + \cdots\right) \end{aligned}$$

Next we need to do a little work on a couple of these terms.

$$\begin{aligned} R_n &= a_{n+1}\left(1 + \frac{a_{n+2}}{a_{n+1}} + \frac{a_{n+3}}{a_{n+1}}\frac{a_{n+2}}{a_{n+2}} + \frac{a_{n+4}}{a_{n+1}}\frac{a_{n+2}}{a_{n+2}}\frac{a_{n+3}}{a_{n+3}} + \cdots\right) \\ &= a_{n+1}\left(1 + \frac{a_{n+2}}{a_{n+1}} + \frac{a_{n+2}}{a_{n+1}}\frac{a_{n+3}}{a_{n+2}} + \frac{a_{n+2}}{a_{n+1}}\frac{a_{n+3}}{a_{n+2}}\frac{a_{n+4}}{a_{n+3}} + \cdots\right) \end{aligned}$$

Now use the definition of r_n to write this as,

$$R_n = a_{n+1}\left(1 + r_{n+1} + r_{n+1}r_{n+2} + r_{n+1}r_{n+2}r_{n+3} + \cdots\right)$$

Okay now let's do the proof.

For the first part we are assuming that $\{r_n\}$ is decreasing and so we can estimate the remainder as,

$$\begin{aligned} R_n &= a_{n+1}\left(1 + r_{n+1} + r_{n+1}r_{n+2} + r_{n+1}r_{n+2}r_{n+3} + \cdots\right) \\ &\le a_{n+1}\left(1 + r_{n+1} + r_{n+1}^2 + r_{n+1}^3 + \cdots\right) \\ &= a_{n+1}\sum_{k=0}^{\infty} r_{n+1}^k \end{aligned}$$

Finally, the series here is a geometric series and because $r_{n+1} < 1$ we know that it converges and we can compute its value. So,

$$R_n \le \frac{a_{n+1}}{1-r_{n+1}}$$

For the second part we are assuming that $\left\{r_n\right\}$ is increasing and we know that,

$$\lim_{n\to\infty}\left|r_n\right| = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = L$$

and so we know that $r_n < L$ for all *n*. The remainder can then be estimated as,

$$\begin{aligned} R_n &= a_{n+1}\left(1+r_{n+1}+r_{n+1}r_{n+2}+r_{n+1}r_{n+2}r_{n+3}+\cdots\right) \\ &\le a_{n+1}\left(1+L+L^2+L^3+\cdots\right) \\ &= a_{n+1}\sum_{k=0}^{\infty}L^k \end{aligned}$$

This is a geometric series and since we are assuming that our original series converges we also know that $L<1$ and so the geometric series above converges and we can compute its value. So,

$$R_n \le \frac{a_{n+1}}{1-L}$$

■

Note that there are some restrictions on the sequence $\left\{r_n\right\}$ and at least one of its terms in order to use these formulas. If the restrictions aren't met then the formulas can't be used.

Let's take a look at an example of this.

Example 4 Using $n=15$ to estimate the value of $\displaystyle\sum_{n=0}^{\infty}\frac{n}{3^n}$.

Solution
First, let's use the ratio test to verify that this is a convergent series.

$$L = \lim_{n\to\infty}\left|\frac{n+1}{3^{n+1}}\,\frac{3^n}{n}\right| = \lim_{n\to\infty}\frac{n+1}{3n} = \frac{1}{3} < 1$$

So, it is convergent. Now let's get the estimate.

$$s_{15} = \sum_{n=0}^{15}\frac{n}{3^n} = 0.7499994250$$

To determine an estimate on the remainder, and hence the error, let's first get the sequence $\left\{r_n\right\}$.

$$r_n = \frac{n+1}{3^{n+1}}\,\frac{3^n}{n} = \frac{n+1}{3n} = \frac{1}{3}\left(1+\frac{1}{n}\right)$$

The last rewrite was just to simplify some of the computations a little. Now, notice that,

$$f\left(x\right) = \frac{1}{3}\left(1+\frac{1}{x}\right) \qquad\qquad f'\left(x\right) = -\frac{1}{3x^2} < 0$$

Since this function is always decreasing and $f\left(n\right) = r_n$ and so, this sequence is decreasing.

Also note that $r_{16} = \frac{1}{3}\left(1+\frac{1}{16}\right) < 1$. Therefore we can use the first case from the fact above to get,

$$R_{15} \le \frac{a_{16}}{1-r_{16}} = \frac{\frac{16}{3^{16}}}{1-\frac{1}{3}\left(1+\frac{1}{16}\right)} = 0.0000005755187$$

So, it looks like our estimate is probably quite good. In this case the exact value is known.

$$\sum_{n=0}^{\infty} \frac{n}{3^n} = \frac{3}{4}$$

and so we can compute the actual error.

$$\left|R_{15}\right| = \left|s - s_{15}\right| = 0.000000575$$

This is less than the upper bound, but unlike in the previous example this actual error is quite close to the upper bound.

In the last two examples we've seen that the upper bound computations on the error can sometimes be quite close to the actual error and at other times they can be off by quite a bit. There is usually no way of knowing ahead of time which it will be and without the exact value in hand there will never be a way of determining which it will be.

Notice that this method did require the series terms to be positive, but that doesn't mean that we can't deal with ratio test series if they have negative terms. Often series that we used ratio test on are also alternating series and so if that is the case we can always resort to the previous material to get an upper bound on the error in the estimation, even if we didn't use the alternating series test to show convergence.

Note however that if the series does have negative terms, but doesn't happen to be an alternating series then we can't use any of the methods discussed in this section to get an upper bound on the error.

Power Series

We've spent quite a bit of time talking about series now and with only a couple of exceptions we've spent most of that time talking about how to determine if a series will converge or not. It's now time to start looking at some specific kinds of series and we'll eventually reach the point where we can talk about a couple of applications of series.

In this section we are going to start talking about power series. A **power series about a**, or just **power series**, is any series that can be written in the form,

$$\sum_{n=0}^{\infty} c_n \left(x-a\right)^n$$

where a and c_n are numbers. The c_n's are often called the **coefficients** of the series. The first thing to notice about a power series is that it is a function of x. That is different from any other kind of series that we've looked at to this point. In all the prior sections we've only allowed numbers in the series and now we are allowing variables to be in the series as well. This will not change how things work however. Everything that we know about series still holds.

In the discussion of power series convergence is still a major question that we'll be dealing with. The difference is that the convergence of the series will now depend upon the values of x that we put into the series. A power series may converge for some values of x and not for other values of x.

Before we get too far into power series there is some terminology that we need to get out of the way.

First, as we will see in our examples, we will be able to show that there is a number R so that the power series will converge for, $\left|x-a\right|<R$ and will diverge for $\left|x-a\right|>R$. This number is called the **radius of convergence** for the series. Note that the series may or may not converge if $\left|x-a\right|=R$. What happens at these points will not change the radius of convergence.

Secondly, the interval of all x's, including the endpoints if need be, for which the power series converges is called the **interval of convergence** of the series.

These two concepts are fairly closely tied together. If we know that the radius of convergence of a power series is R then we have the following.

$$\begin{array}{ll} a-R<x<a+R & \text{power series converges} \\ x<a-R \ \text{ and } \ x>a+R & \text{power series diverges} \end{array}$$

The interval of convergence must then contain the interval $a-R<x<a+R$ since we know that the power series will converge for these values. We also know that the interval of convergence can't contain x's in the ranges $x<a-R$ and $x>a+R$ since we know the power series diverges for these value of x. Therefore, to completely identify the interval of convergence all that we have to do is determine if the power series will converge for $x=a-R$ or $x=a+R$. If the power series converges for one or both of these values then we'll need to include those in the interval of convergence.

Before getting into some examples let's take a quick look at the convergence of a power series for the case of $x = a$. In this case the power series becomes,

$$\sum_{n=0}^{\infty} c_n \left(a-a\right)^n = \sum_{n=0}^{\infty} c_n \left(0\right)^n = c_0 \left(0\right)^0 + \sum_{n=1}^{\infty} c_n \left(0\right)^n = c_0 + \sum_{n=1}^{\infty} 0 = c_0 + 0 = c_0$$

and so the power series converges. Note that we had to strip out the first term since it was the only non-zero term in the series.

It is important to note that no matter what else is happening in the power series we are guaranteed to get convergence for $x = a$. The series may not converge for any other value of x, but it will always converge for $x = a$.

Let's work some examples. We'll put quite a bit of detail into the first example and then not put quite as much detail in the remaining examples.

Example 1 Determine the radius of convergence and interval of convergence for the following power series.

$$\sum_{n=1}^{\infty} \frac{(-1)^n n}{4^n} \left(x+3\right)^n$$

Solution

Okay, we know that this power series will converge for $x = -3$, but that's it at this point. To determine the remainder of the x's for which we'll get convergence we can use any of the tests that we've discussed to this point. After application of the test that we choose to work with we will arrive at condition(s) on x that we can use to determine which values of x for which the power series will converge and which values of x for which the power series will diverge. From this we can get the radius of convergence and most of the interval of convergence (with the possible exception of the endpoints).

With all that said, the best tests to use here are almost always the ratio or root test. Most of the power series that we'll be looking at are set up for one or the other. In this case we'll use the ratio test.

$$\begin{aligned} L &= \lim_{n\to\infty} \left| \frac{(-1)^{n+1} \left(n+1\right)\left(x+3\right)^{n+1}}{4^{n+1}} \frac{4^n}{(-1)^n \left(n\right)\left(x+3\right)^n} \right| \\ &= \lim_{n\to\infty} \left| \frac{-\left(n+1\right)\left(x+3\right)}{4n} \right| \end{aligned}$$

Before going any farther with the limit let's notice that since x is not dependent on the limit it can be factored out of the limit. Notice as well that in doing this we'll need to keep the absolute value bars on it since we need to make sure everything stays positive and x could well be a value that will make things negative. The limit is then,

$$\begin{aligned} L &= \left|x+3\right| \lim_{n\to\infty} \frac{n+1}{4n} \\ &= \frac{1}{4}\left|x+3\right| \end{aligned}$$

So, the ratio test tells us that if $L < 1$ the series will converge, if $L > 1$ the series will diverge, and if $L = 1$ we don't know what will happen. So, we have,

$$\frac{1}{4}\left|x+3\right| < 1 \quad \Rightarrow \quad \left|x+3\right| < 4 \qquad \text{series converges}$$

$$\frac{1}{4}\left|x+3\right| > 1 \quad \Rightarrow \quad \left|x+3\right| > 4 \qquad \text{series diverges}$$

We'll deal with the $L = 1$ case in a bit. Notice that we now have the radius of convergence for this power series. These are exactly the conditions required for the radius of convergence. The radius of convergence for this power series is $R = 4$.

Now, let's get the interval of convergence. We'll get most (if not all) of the interval by solving the first inequality from above.

$$-4 < x+3 < 4$$
$$-7 < x < 1$$

So, most of the interval of validity is given by $-7 < x < 1$. All we need to do is determine if the power series will converge or diverge at the endpoints of this interval. Note that these values of *x* will correspond to the value of *x* that will give $L = 1$.

The way to determine convergence at these points is to simply plug them into the original power series and see if the series converges or diverges using any test necessary.

$x = -7$:
In this case the series is,

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{(-1)^n n}{4^n}(-4)^n &= \sum_{n=1}^{\infty}\frac{(-1)^n n}{4^n}(-1)^n 4^n \\ &= \sum_{n=1}^{\infty}(-1)^n(-1)^n n \qquad\qquad (-1)^n(-1)^n = (-1)^{2n} = 1 \\ &= \sum_{n=1}^{\infty} n\end{aligned}$$

This series is divergent by the Divergence Test since $\lim\limits_{n\to\infty} n = \infty \neq 0$.

$x = 1$:
In this case the series is,

$$\sum_{n=1}^{\infty}\frac{(-1)^n n}{4^n}(4)^n = \sum_{n=1}^{\infty}(-1)^n n$$

This series is also divergent by the Divergence Test since $\lim\limits_{n\to\infty}(-1)^n n$ doesn't exist.

So, in this case the power series will not converge for either endpoint. The interval of convergence is then,

$$-7 < x < 1$$

In the previous example the power series didn't converge for either endpoint of the interval. Sometimes that will happen, but don't always expect that to happen. The power series could converge at either both of the endpoints or only one of the endpoints.

Example 2 Determine the radius of convergence and interval of convergence for the following power series.

$$\sum_{n=1}^{\infty} \frac{2^n}{n}(4x-8)^n$$

Solution

Let's jump right into the ratio test.

$$\begin{aligned} L &= \lim_{n\to\infty}\left|\frac{2^{n+1}(4x-8)^{n+1}}{n+1}\,\frac{n}{2^n(4x-8)^n}\right| \\ &= \lim_{n\to\infty}\left|\frac{2n(4x-8)}{n+1}\right| \\ &= |4x-8|\lim_{n\to\infty}\frac{2n}{n+1} \\ &= 2|4x-8| \end{aligned}$$

So we will get the following convergence/divergence information from this.

$$\begin{aligned} 2|4x-8| &< 1 \qquad \text{series converges} \\ 2|4x-8| &> 1 \qquad \text{series diverges} \end{aligned}$$

We need to be careful here in determining the interval of convergence. The interval of convergence requires $|x-a|<R$ and $|x-a|>R$. In other words, we need to factor a 4 out of the absolute value bars in order to get the correct radius of convergence. Doing this gives,

$$\begin{aligned} 8|x-2| &< 1 \qquad \Rightarrow \qquad |x-2| < \frac{1}{8} \qquad \text{series converges} \\ 8|x-2| &> 1 \qquad \Rightarrow \qquad |x-2| > \frac{1}{8} \qquad \text{series diverges} \end{aligned}$$

So, the radius of convergence for this power series is $R=\frac{1}{8}$.

Now, let's find the interval of convergence. Again, we'll first solve the inequality that gives convergence above.

$$\begin{gathered} -\frac{1}{8} < x-2 < \frac{1}{8} \\ \frac{15}{8} < x < \frac{17}{8} \end{gathered}$$

Now check the endpoints.

$x = \frac{15}{8}$:

The series here is,

$$\sum_{n=1}^{\infty} \frac{2^n}{n}\left(\frac{15}{2} - 8\right)^n = \sum_{n=1}^{\infty} \frac{2^n}{n}\left(-\frac{1}{2}\right)^n$$
$$= \sum_{n=1}^{\infty} \frac{2^n}{n} \frac{(-1)^n}{2^n}$$
$$= \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

This is the alternating harmonic series and we know that it converges.

$x = \frac{17}{8}$:

The series here is,

$$\sum_{n=1}^{\infty} \frac{2^n}{n}\left(\frac{17}{2} - 8\right)^n = \sum_{n=1}^{\infty} \frac{2^n}{n}\left(\frac{1}{2}\right)^n$$
$$= \sum_{n=1}^{\infty} \frac{2^n}{n} \frac{1}{2^n}$$
$$= \sum_{n=1}^{\infty} \frac{1}{n}$$

This is the harmonic series and we know that it diverges.

So, the power series converges for one of the endpoints, but not the other. This will often happen so don't get excited about it when it does. The interval of convergence for this power series is then,

$$\frac{15}{8} \le x < \frac{17}{8}$$

We now need to take a look at a couple of special cases with radius and intervals of convergence.

Example 3 Determine the radius of convergence and interval of convergence for the following power series.

$$\sum_{n=0}^{\infty} n!(2x+1)^n$$

Solution

We'll start this example with the ratio test as we have for the previous ones.

$$\begin{aligned} L &= \lim_{n\to\infty}\left|\frac{(n+1)!(2x+1)^{n+1}}{n!(2x+1)^n}\right| \\ &= \lim_{n\to\infty}\left|\frac{(n+1)n!\,(2x+1)}{n!}\right| \\ &= |2x+1|\lim_{n\to\infty}(n+1) \end{aligned}$$

At this point we need to be careful. The limit is infinite, but there is that term with the *x*'s in front of the limit. We'll have $L = \infty > 1$ provided $x \neq -\frac{1}{2}$.

So, this power series will only converge if $x = -\frac{1}{2}$. If you think about it we actually already knew that however. From our initial discussion we know that every power series will converge for $x = a$ and in this case $a = -\frac{1}{2}$. Remember that we get *a* from $(x-a)^n$, and notice the coefficient of the *x* must be a one!

In this case we say the radius of convergence is $R = 0$ and the interval of convergence is $x = -\frac{1}{2}$, and yes we really did mean interval of convergence even though it's only a point.

Example 4 Determine the radius of convergence and interval of convergence for the following power series.

$$\sum_{n=1}^{\infty}\frac{(x-6)^n}{n^n}$$

Solution

In this example the root test seems more appropriate. So,

$$\begin{aligned} L &= \lim_{n\to\infty}\left|\frac{(x-6)^n}{n^n}\right|^{\frac{1}{n}} \\ &= \lim_{n\to\infty}\left|\frac{x-6}{n}\right| \\ &= |x-6|\lim_{n\to\infty}\frac{1}{n} \\ &= 0 \end{aligned}$$

So, since $L = 0 < 1$ regardless of the value of *x* this power series will converge for every *x*.

In these cases we say that the radius of convergence is $R = \infty$ and interval of convergence is $-\infty < x < \infty$.

So, let's summarize the last two examples. If the power series only converges for $x = a$ then the radius of convergence is $R = 0$ and the interval of convergence is $x = a$. Likewise if the power series converges for every *x* the radius of convergence is $R = \infty$ and interval of convergence is $-\infty < x < \infty$.

Let’s work one more example.

Example 5 Determine the radius of convergence and interval of convergence for the following power series.

$$\sum_{n=1}^{\infty} \frac{x^{2n}}{(-3)^n}$$

Solution

First notice that $a = 0$ in this problem. That’s not really important to the problem, but it’s worth pointing out so people don’t get excited about it.

The important difference in this problem is the exponent on the *x*. In this case it is 2*n* rather than the standard *n*. As we will see some power series will have exponents other than an *n* and so we still need to be able to deal with these kinds of problems.

This one seems set up for the root test again so let’s use that.

$$\begin{aligned} L &= \lim_{n\to\infty} \left| \frac{x^{2n}}{(-3)^n} \right|^{\frac{1}{n}} \\ &= \lim_{n\to\infty} \left| \frac{x^2}{-3} \right| \\ &= \frac{x^2}{3} \end{aligned}$$

So, we will get convergence if

$$\frac{x^2}{3} < 1 \qquad \Rightarrow \qquad x^2 < 3$$

The radius of convergence is NOT 3 however. The radius of convergence requires an exponent of 1 on the *x*. Therefore,

$$\begin{aligned} \sqrt{x^2} &< \sqrt{3} \\ |x| &< \sqrt{3} \end{aligned}$$

Be careful with the absolute value bars! In this case it looks like the radius of convergence is $R = \sqrt{3}$. Notice that we didn’t bother to put down the inequality for divergence this time. The inequality for divergence is just the interval for convergence that the test gives with the inequality switched and generally isn’t needed. We will usually skip that part.

Now let’s get the interval of convergence. First from the inequality we get,

$$-\sqrt{3} < x < \sqrt{3}$$

Now check the endpoints.

$x = -\sqrt{3}$:

Here the power series is,

$$\sum_{n=1}^{\infty}\frac{\left(-\sqrt{3}\right)^{2n}}{\left(-3\right)^{n}}=\sum_{n=1}^{\infty}\frac{\left(\left(-\sqrt{3}\right)^{2}\right)^{n}}{\left(-3\right)^{n}}$$

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

$$=\sum_{n=1}^{\infty}\left(-1\right)^{n}$$

This series is divergent by the Divergence Test since $\lim_{n\to\infty}\left(-1\right)^{n}$ doesn't exist.

$x=\sqrt{3}$:
Because we're squaring the *x* this series will be the same as the previous step.

$$\sum_{n=1}^{\infty}\frac{\left(\sqrt{3}\right)^{2n}}{\left(-3\right)^{n}}=\sum_{n=1}^{\infty}\left(-1\right)^{n}$$

which is divergent.

The interval of convergence is then,

$$-\sqrt{3}<x<\sqrt{3}$$

Power Series and Functions

We opened the last section by saying that we were going to start thinking about applications of series and then promptly spent the section talking about convergence again. It's now time to actually start with the applications of series.

With this section we will start talking about how to represent functions with power series. The natural question of why we might want to do this will be answered in a couple of sections once we actually learn how to do this.

Let's start off with one that we already know how to do, although when we first ran across this series we didn't think of it as a power series nor did we acknowledge that it represented a function.

Recall that the geometric series is

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r} \qquad \text{provided } \left|r\right| < 1$$

Don't forget as well that if $\left|r\right| \geq 1$ the series diverges.

Now, if we take $a = 1$ and $r = x$ this becomes,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \qquad \text{provided } \left|x\right| < 1 \tag{1}$$

Turning this around we can see that we can represent the function

$$f\left(x\right) = \frac{1}{1-x} \tag{2}$$

with the power series

$$\sum_{n=0}^{\infty} x^n \qquad \text{provided } \left|x\right| < 1 \tag{3}$$

This provision is important. We can clearly plug any number other than $x = 1$ into the function, however, we will only get a convergent power series if $\left|x\right| < 1$. This means the equality in (1) will only hold if $\left|x\right| < 1$. For any other value of *x* the equality won't hold. Note as well that we can also use this to acknowledge that the radius of convergence of this power series is $R = 1$ and the interval of convergence is $\left|x\right| < 1$.

This idea of convergence is important here. We will be representing many functions as power series and it will be important to recognize that the representations will often only be valid for a range of *x*'s and that there may be values of *x* that we can plug into the function that we can't plug into the power series representation.

In this section we are going to concentrate on representing functions with power series where the functions can be related back to (2).

In this way we will hopefully become familiar with some of the kinds of manipulations that we will sometimes need to do when working with power series.

So, let's jump into a couple of examples.

Example 1 Find a power series representation for the following function and determine its interval of convergence.

$$g(x) = \frac{1}{1+x^3}$$

Solution

What we need to do here is to relate this function back to (2). This is actually easier than it might look. Recall that the *x* in (2) is simply a variable and can represent anything. So, a quick rewrite of $g(x)$ gives,

$$g(x) = \frac{1}{1-\left(-x^3\right)}$$

and so the $-x^3$ in $g(x)$ holds the same place as the *x* in (2). Therefore, all we need to do is replace the *x* in (3) and we've got a power series representation for $g(x)$.

$$g(x) = \sum_{n=0}^{\infty}\left(-x^3\right)^n \qquad\qquad \text{provided } \left|-x^3\right| < 1$$

Notice that we replaced both the *x* in the power series and in the interval of convergence.

All we need to do now is a little simplification.

$$g(x) = \sum_{n=0}^{\infty}(-1)^n x^{3n} \qquad\qquad \text{provided } |x|^3 < 1 \qquad \Rightarrow \qquad |x| < 1$$

So, in this case the interval of convergence is the same as the original power series. This usually won't happen. More often than not the new interval of convergence will be different from the original interval of convergence.

Example 2 Find a power series representation for the following function and determine its interval of convergence.

$$h(x) = \frac{2x^2}{1+x^3}$$

Solution

This function is similar to the previous function. The difference is the numerator and at first glance that looks to be an important difference. Since (2) doesn't have an *x* in the numerator it appears that we can't relate this function back to that.

However, now that we've worked the first example this one is actually very simple since we can use the result of the answer from that example. To see how to do this let's first rewrite the function a little.

$$h(x) = 2x^2\,\frac{1}{1+x^3}$$

Now, from the first example we've already got a power series for the second term so let's use that to write the function as,

$$h(x) = 2x^2 \sum_{n=0}^{\infty} (-1)^n x^{3n} \qquad\qquad \text{provided } |x| < 1$$

Notice that the presence of x's outside of the series will NOT affect its convergence and so the interval of convergence remains the same.

The last step is to bring the coefficient into the series and we'll be done. When we do this make sure and combine the x's as well. We typically only want a single x in a power series.

$$h(x) = \sum_{n=0}^{\infty} 2(-1)^n x^{3n+2} \qquad \text{provided } |x| < 1$$

As we saw in the previous example we can often use previous results to help us out. This is an important idea to remember as it can often greatly simplify our work.

Example 3 Find a power series representation for the following function and determine its interval of convergence.

$$f(x) = \frac{x}{5-x}$$

Solution

So, again, we've got an x in the numerator. So, as with the last example let's factor that out and see what we've got left.

$$f(x) = x\,\frac{1}{5-x}$$

If we had a power series representation for

$$g(x) = \frac{1}{5-x}$$

we could get a power series representation for $f(x)$.

So, let's find one. We'll first notice that in order to use (4) we'll need the number in the denominator to be a one. That's easy enough to get.

$$g(x) = \frac{1}{5}\,\frac{1}{1-\frac{x}{5}}$$

Now all we need to do to get a power series representation is to replace the x in (3) with $\frac{x}{5}$. Doing this gives,

$$g(x) = \frac{1}{5}\sum_{n=0}^{\infty}\left(\frac{x}{5}\right)^n \qquad\qquad \text{provided } \left|\frac{x}{5}\right| < 1$$

Now let's do a little simplification on the series.

$$
\begin{aligned}
g(x) &= \frac{1}{5}\sum_{n=0}^{\infty}\frac{x^n}{5^n} \\
&= \sum_{n=0}^{\infty}\frac{x^n}{5^{n+1}}
\end{aligned}
$$

The interval of convergence for this series is,

$$\left|\frac{x}{5}\right| < 1 \qquad \Rightarrow \qquad \frac{1}{5}|x| < 1 \qquad \Rightarrow \qquad |x| < 5$$

Okay, this was the work for the power series representation for $g(x)$ let's now find a power series representation for the original function. All we need to do for this is to multiply the power series representation for $g(x)$ by x and we'll have it.

$$
\begin{aligned}
f(x) &= x\,\frac{1}{5-x} \\
&= x\sum_{n=0}^{\infty}\frac{x^n}{5^{n+1}} \\
&= \sum_{n=0}^{\infty}\frac{x^{n+1}}{5^{n+1}}
\end{aligned}
$$

The interval of convergence doesn't change and so it will be $|x| < 5$.

So, hopefully we now have an idea on how to find the power series representation for some functions. Admittedly all of the functions could be related back to (2) but it's a start.

We now need to look at some further manipulation of power series that we will need to do on occasion. We need to discuss differentiation and integration of power series.

Let's start with differentiation of the power series,

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots$$

Now, we know that if we differentiate a finite sum of terms all we need to do is differentiate each of the terms and then add them back up. With infinite sums there are some subtleties involved that we need to be careful with, but are somewhat beyond the scope of this course.

Nicely enough for us however, it is known that if the power series representation of $f(x)$ has a radius of convergence of $R > 0$ then the term by term differentiation of the power series will also have a radius of convergence of R and (more importantly) will in fact be the power series representation of $f'(x)$ provided we stay within the radius of convergence.

Again, we should make the point that if we aren't dealing with a power series then we may or may not be able to differentiate each term of the series to get the derivative of the series.

So, what all this means for us is that,

$$f'(x) = \frac{d}{dx}\sum_{n=0}^{\infty} c_n (x-a)^n = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \cdots = \sum_{n=1}^{\infty} n c_n (x-a)^{n-1}$$

Note the initial value of this series. It has been changed from $n = 0$ to $n = 1$. This is an acknowledgement of the fact that the derivative of the first term is zero and hence isn't in the derivative. Notice however, that since the *n=0* term of the above series is also zero, we could start the series at $n = 0$ if it was required for a particular problem. In general however, this won't be done in this class.

We can now find formulas for higher order derivatives as well now.

$$\begin{aligned} f''(x) &= \sum_{n=2}^{\infty} n(n-1)c_n (x-a)^{n-2} \\ f'''(x) &= \sum_{n=3}^{\infty} n(n-1)(n-2)c_n (x-a)^{n-3} \\ &etc. \end{aligned}$$

Once again, notice that the initial value of *n* changes with each differentiation in order to acknowledge that a term from the original series differentiated to zero.

Let's now briefly talk about integration. Just as with the differentiation, when we've got an infinite series we need to be careful about just integration term by term. Much like with derivatives it turns out that as long as we're working with power series we can just integrate the terms of the series to get the integral of the series itself. In other words,

$$\begin{aligned} \int f(x)\,dx &= \int \sum_{n=0}^{\infty} c_n (x-a)^n \, dx \\ &= \sum_{n=0}^{\infty} \int c_n (x-a)^n \, dx \\ &= C + \sum_{n=0}^{\infty} c_n \frac{(x-a)^{n+1}}{n+1} \end{aligned}$$

Notice that we pick up a constant of integration, *C*, that is outside the series here.

Let's summarize the differentiation and integration ideas before moving on to an example or two.

Fact

If $f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$ has a radius of convergence of $R > 0$ then,

$$f'(x) = \sum_{n=1}^{\infty} n c_n (x-a)^{n-1} \qquad\qquad \int f(x)\,dx = C + \sum_{n=0}^{\infty} c_n \frac{(x-a)^{n+1}}{n+1}$$

and both of these also have a radius of convergence of *R*.

Now, let's see how we can use these facts to generate some more power series representations of functions.

Example 4 Find a power series representation for the following function and determine its interval of convergence.

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

Solution

To do this problem let's notice that

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

Then since we've got a power series representation for

$$\frac{1}{1-x}$$

all that we'll need to do is differentiate that power series to get a power series representation for $g(x)$.

$$\begin{aligned} g(x) &= \frac{1}{(1-x)^2} \\ &= \frac{d}{dx}\left(\frac{1}{1-x}\right) \\ &= \frac{d}{dx}\left(\sum_{n=0}^{\infty} x^n\right) \\ &= \sum_{n=1}^{\infty} n x^{n-1} \end{aligned}$$

Then since the original power series had a radius of convergence of $R = 1$ the derivative, and hence *g(x)*, will also have a radius of convergence of $R = 1$.

Example 5 Find a power series representation for the following function and determine its interval of convergence.

$$h(x) = \ln(5-x)$$

Solution

In this case we need to notice that

$$\int \frac{1}{5-x}\,dx = -\ln(5-x)$$

and the recall that we have a power series representation for

$$\frac{1}{5-x}$$

Remember we found a representation for this in Example 3. So,

$$\begin{aligned}\ln(5-x) &= -\int \frac{1}{5-x}\,dx \\ &= -\int \sum_{n=0}^{\infty} \frac{x^n}{5^{n+1}}\,dx \\ &= C - \sum_{n=0}^{\infty} \frac{x^{n+1}}{(n+1)5^{n+1}}\end{aligned}$$

We can find the constant of integration, *C*, by plugging in a value of *x*. A good choice is $x = 0$ since that will make the series easy to evaluate.

$$\begin{aligned}\ln(5-0) &= C - \sum_{n=0}^{\infty} \frac{0^{n+1}}{(n+1)5^{n+1}} \\ \ln(5) &= C\end{aligned}$$

So, the final answer is,

$$\ln(5-x) = \ln(5) - \sum_{n=0}^{\infty} \frac{x^{n+1}}{(n+1)5^{n+1}}$$

Note that it is okay to have the constant sitting outside of the series like this. In fact, there is no way to bring it into the series so don't get excited about it.

Taylor Series

In the previous section we started looking at writing down a power series representation of a function. The problem with the approach in that section is that everything came down to needing to be able to relate the function in some way to

$$\frac{1}{1-x}$$

and while there are many functions out there that can be related to this function there are many more that simply can't be related to this.

So, without taking anything away from the process we looked at in the previous section, what we need to do is come up with a more general method for writing a power series representation for a function.

So, for the time being, let's make two assumptions. First, let's assume that the function $f(x)$ does in fact have a power series representation about $x=a$,

$$f(x)=\sum_{n=0}^{\infty}c_n(x-a)^n=c_0+c_1(x-a)+c_2(x-a)^2+c_3(x-a)^3+c_4(x-a)^4+\cdots$$

Next, we will need to assume that the function, $f(x)$, has derivatives of every order and that we can in fact find them all.

Now that we've assumed that a power series representation exists we need to determine what the coefficients, c_n, are. This is easier than it might at first appear to be. Let's first just evaluate everything at $x=a$. This gives,

$$f(a)=c_0$$

So, all the terms except the first are zero and we now know what c_0 is. Unfortunately, there isn't any other value of x that we can plug into the function that will allow us to quickly find any of the other coefficients. However, if we take the derivative of the function (and its power series) then plug in $x=a$ we get,

$$\begin{aligned} f'(x)&=c_1+2c_2(x-a)+3c_3(x-a)^2+4c_4(x-a)^3+\cdots \\ f'(a)&=c_1 \end{aligned}$$

and we now know c_1.

Let's continue with this idea and find the second derivative.

$$\begin{aligned} f''(x)&=2c_2+3(2)c_3(x-a)+4(3)c_4(x-a)^2+\cdots \\ f''(a)&=2c_2 \end{aligned}$$

So, it looks like,

$$c_2=\frac{f''(a)}{2}$$

Using the third derivative gives,

$$f'''(x) = 3(2)c_3 + 4(3)(2)c_4(x-a) + \cdots$$
$$f'''(a) = 3(2)c_3 \qquad \Rightarrow \qquad c_3 = \frac{f'''(a)}{3(2)}$$

Using the fourth derivative gives,

$$f^{(4)}(x) = 4(3)(2)c_4 + 5(4)(3)(2)c_5(x-a)\cdots$$
$$f^{(4)}(a) = 4(3)(2)c_4 \qquad \Rightarrow \qquad c_4 = \frac{f^{(4)}(a)}{4(3)(2)}$$

Hopefully by this time you've seen the pattern here. It looks like, in general, we've got the following formula for the coefficients.

$$c_n = \frac{f^{(n)}(a)}{n!}$$

This even works for *n=0* if you recall that $0! = 1$ and define $f^{(0)}(x) = f(x)$.

So, provided a power series representation for the function $f(x)$ about $x = a$ exists the **Taylor Series for** $f(x)$ **about** $x = a$ is,

Taylor Series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$
$$= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

If we use $a = 0$, so we are talking about the Taylor Series about $x = 0$, we call the series a **Maclaurin Series** for $f(x)$ or,

Maclaurin Series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n$$
$$= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots$$

Before working any examples of Taylor Series we first need to address the assumption that a Taylor Series will in fact exist for a given function. Let's start out with some notation and definitions that we'll need.

To determine a condition that must be true in order for a Taylor series to exist for a function let's first define the **nth degree Taylor polynomial of** $f(x)$ as,

$$T_n(x) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i$$

Note that this really is a polynomial of degree at most *n*. If we were to write out the sum without the summation notation this would clearly be an n^{th} degree polynomial. We'll see a nice application of Taylor polynomials in the next section.

Notice as well that for the full Taylor Series,

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

the n^{th} degree Taylor polynomial is just the partial sum for the series.

Next, the **remainder** is defined to be,

$$R_n(x) = f(x) - T_n(x)$$

So, the remainder is really just the *error* between the function $f(x)$ and the n^{th} degree Taylor polynomial for a given *n*.

With this definition note that we can then write the function as,

$$f(x) = T_n(x) + R_n(x)$$

We now have the following Theorem.

Theorem

Suppose that $f(x) = T_n(x) + R_n(x)$. Then if,

$$\lim_{n\to\infty} R_n(x) = 0$$

for $|x-a| < R$ then,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

on $|x-a| < R$.

In general showing that $\lim_{n\to\infty} R_n(x) = 0$ is a somewhat difficult process and so we will be assuming that this can be done for some *R* in all of the examples that we'll be looking at.

Now let's look at some examples.

Example 1 Find the Taylor Series for $f(x) = \mathbf{e}^x$ about $x = 0$.

Solution

This is actually one of the easier Taylor Series that we'll be asked to compute. To find the Taylor Series for a function we will need to determine a general formula for $f^{(n)}(a)$. This is one of the few functions where this is easy to do right from the start.

To get a formula for $f^{(n)}(0)$ all we need to do is recognize that,

$$f^{(n)}(x) = \mathbf{e}^x \qquad\qquad n = 0,1,2,3,\ldots$$

and so,

$$f^{(n)}(0) = \mathbf{e}^0 = 1 \qquad\qquad n = 0,1,2,3,\ldots$$

Therefore, the Taylor series for $f(x) = \mathbf{e}^x$ about *x=0* is,

$$\mathbf{e}^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Example 2 Find the Taylor Series for $f(x) = \mathbf{e}^{-x}$ about $x = 0$.

Solution

There are two ways to do this problem. Both are fairly simple, however one of them requires significantly less work. We'll work both solutions since the longer one has some nice ideas that we'll see in other examples.

Solution 1

As with the first example we'll need to get a formula for $f^{(n)}(0)$. However, unlike the first one we've got a little more work to do. Let's first take some derivatives and evaluate them at *x=0*.

$$\begin{aligned}
f^{(0)}(x) &= \mathbf{e}^{-x} & f^{(0)}(0) &= 1 \\
f^{(1)}(x) &= -\mathbf{e}^{-x} & f^{(1)}(0) &= -1 \\
f^{(2)}(x) &= \mathbf{e}^{-x} & f^{(2)}(0) &= 1 \\
f^{(3)}(x) &= -\mathbf{e}^{-x} & f^{(3)}(0) &= -1 \\
&\vdots & &\vdots \\
f^{(n)}(x) &= (-1)^n \mathbf{e}^{-x} & f^{(n)}(0) &= (-1)^n \qquad n = 0,1,2,3
\end{aligned}$$

After a couple of computations we were able to get general formulas for both $f^{(n)}(x)$ and $f^{(n)}(0)$. We often won't be able to get a general formula for $f^{(n)}(x)$ so don't get too excited about getting that formula. Also, as we will see it won't always be easy to get a general formula for $f^{(n)}(a)$.

So, in this case we've got general formulas so all we need to do is plug these into the Taylor Series formula and be done with the problem.

$$\mathbf{e}^{-x} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!}$$

Solution 2
The previous solution wasn't too bad and we often have to do things in that manner. However, in this case there is a much shorter solution method. In the previous section we used series that we've already found to help us find a new series. Let's do the same thing with this one. We already know a Taylor Series for $\mathbf{e}^{x}$ about $x = 0$ and in this case the only difference is we've got a "-*x*" in the exponent instead of just an *x*.

So, all we need to do is replace the *x* in the Taylor Series that we found in the first example with "-*x*".

$$\mathbf{e}^{-x} = \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!}$$

This is a much shorter method of arriving at the same answer so don't forget about using previously computed series where possible (and allowed of course).

Example 3 Find the Taylor Series for $f(x) = x^4 \mathbf{e}^{-3x^2}$ about $x = 0$.

Solution
For this example we will take advantage of the fact that we already have a Taylor Series for $\mathbf{e}^{x}$ about $x = 0$. In this example, unlike the previous example, doing this directly would be significantly longer and more difficult.

$$\begin{aligned} x^4 \mathbf{e}^{-3x^2} &= x^4 \sum_{n=0}^{\infty} \frac{\left(-3x^2\right)^n}{n!} \\ &= x^4 \sum_{n=0}^{\infty} \frac{(-3)^n x^{2n}}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(-3)^n x^{2n+4}}{n!} \end{aligned}$$

To this point we've only looked at Taylor Series about $x = 0$ (also known as Maclaurin Series) so let's take a look at a Taylor Series that isn't about $x = 0$. Also, we'll pick on the exponential function one more time since it makes some of the work easier. This will be the final Taylor Series for exponentials in this section.

Example 4 Find the Taylor Series for $f(x) = \mathbf{e}^{-x}$ about $x = -4$.

Solution
Finding a general formula for $f^{(n)}(-4)$ is fairly simple.

$$f^{(n)}(x) = (-1)^n \mathbf{e}^{-x} \qquad\qquad f^{(n)}(-4) = (-1)^n \mathbf{e}^{4}$$

The Taylor Series is then,

$$\mathbf{e}^{-x} = \sum_{n=0}^{\infty} \frac{(-1)^n \mathbf{e}^4}{n!} (x+4)^n$$

Okay, we now need to work some examples that don't involve the exponential function since these will tend to require a little more work.

Example 5 Find the Taylor Series for $f(x) = \cos(x)$ about $x = 0$.

Solution
First we'll need to take some derivatives of the function and evaluate them at *x=0*.

$$\begin{aligned}
f^{(0)}(x) &= \cos x & \qquad f^{(0)}(0) &= 1 \\
f^{(1)}(x) &= -\sin x & \qquad f^{(1)}(0) &= 0 \\
f^{(2)}(x) &= -\cos x & \qquad f^{(2)}(0) &= -1 \\
f^{(3)}(x) &= \sin x & \qquad f^{(3)}(0) &= 0 \\
f^{(4)}(x) &= \cos x & \qquad f^{(4)}(0) &= 1 \\
f^{(5)}(x) &= -\sin x & \qquad f^{(5)}(0) &= 0 \\
f^{(6)}(x) &= -\cos x & \qquad f^{(6)}(0) &= -1 \\
&\vdots & &\vdots
\end{aligned}$$

In this example, unlike the previous ones, there is not an easy formula for either the general derivative or the evaluation of the derivative. However, there is a clear pattern to the evaluations. So, let's plug what we've got into the Taylor series and see what we get,

$$\begin{aligned}
\cos x &= \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \\
&= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \frac{f^{(5)}(0)}{5!}x^5 + \cdots \\
&= \underbrace{1}_{n=0} + \underbrace{0}_{n=1} \underbrace{-\frac{1}{2!}x^2}_{n=2} + \underbrace{0}_{n=3} + \underbrace{\frac{1}{4!}x^4}_{n=4} + \underbrace{0}_{n=5} \underbrace{-\frac{1}{6!}x^6}_{n=6} + \cdots
\end{aligned}$$

So, we only pick up terms with even powers on the x's. This doesn't really help us to get a general formula for the Taylor Series. However, let's drop the zeroes and "renumber" the terms as follows to see what we can get.

$$\cos x = \underbrace{1}_{n=0} \underbrace{-\frac{1}{2!}x^2}_{n=1} + \underbrace{\frac{1}{4!}x^4}_{n=2} \underbrace{-\frac{1}{6!}x^6}_{n=3} + \cdots$$

By renumbering the terms as we did we can actually come up with a general formula for the Taylor Series and here it is,

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

This idea of renumbering the series terms as we did in the previous example isn't used all that often, but occasionally is very useful. There is one more series where we need to do it so let's take a look at that so we can get one more example down of renumbering series terms.

Example 6 Find the Taylor Series for $f(x) = \sin(x)$ about $x = 0$.

Solution

As with the last example we'll start off in the same manner.

$$\begin{aligned}
f^{(0)}(x) &= \sin x & \qquad f^{(0)}(0) &= 0 \\
f^{(1)}(x) &= \cos x & \qquad f^{(1)}(0) &= 1 \\
f^{(2)}(x) &= -\sin x & \qquad f^{(2)}(0) &= 0 \\
f^{(3)}(x) &= -\cos x & \qquad f^{(3)}(0) &= -1 \\
f^{(4)}(x) &= \sin x & \qquad f^{(4)}(0) &= 0 \\
f^{(5)}(x) &= \cos x & \qquad f^{(5)}(0) &= 1 \\
f^{(6)}(x) &= -\sin x & \qquad f^{(6)}(0) &= 0 \\
&\vdots & &\vdots
\end{aligned}$$

So, we get a similar pattern for this one. Let's plug the numbers into the Taylor Series.

$$\begin{aligned}
\sin x &= \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \\
&= \frac{1}{1!}x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \cdots
\end{aligned}$$

In this case we only get terms that have an odd exponent on x and as with the last problem once we ignore the zero terms there is a clear pattern and formula. So renumbering the terms as we did in the previous example we get the following Taylor Series.

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$$

We really need to work another example or two in which $f(x)$ isn't about $x = 0$.

Example 7 Find the Taylor Series for $f(x) = \ln(x)$ about $x = 2$.

Solution

Here are the first few derivatives and the evaluations.

$$
\begin{aligned}
f^{(0)}(x) &= \ln(x) & f^{(0)}(2) &= \ln 2 \\
f^{(1)}(x) &= \frac{1}{x} & f^{(1)}(2) &= \frac{1}{2} \\
f^{(2)}(x) &= -\frac{1}{x^2} & f^{(2)}(2) &= -\frac{1}{2^2} \\
f^{(3)}(x) &= \frac{2}{x^3} & f^{(3)}(2) &= \frac{2}{2^3} \\
f^{(4)}(x) &= -\frac{2(3)}{x^4} & f^{(4)}(2) &= -\frac{2(3)}{2^4} \\
f^{(5)}(x) &= \frac{2(3)(4)}{x^5} & f^{(5)}(2) &= \frac{2(3)(4)}{2^5} \\
&\vdots & &\vdots \\
f^{(n)}(x) &= \frac{(-1)^{n+1}(n-1)!}{x^n} & f^{(n)}(2) &= \frac{(-1)^{n+1}(n-1)!}{2^n} \qquad n = 1, 2, 3, \ldots
\end{aligned}
$$

Note that while we got a general formula here it doesn't work for $n = 0$. This will happen on occasion so don't worry about it when it does.

In order to plug this into the Taylor Series formula we'll need to strip out the $n = 0$ term first.

$$
\begin{aligned}
\ln(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(2)}{n!}(x-2)^n \\
&= f(2) + \sum_{n=1}^{\infty} \frac{f^{(n)}(2)}{n!}(x-2)^n \\
&= \ln(2) + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}(n-1)!}{n!\ 2^n}(x-2)^n \\
&= \ln(2) + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n\,2^n}(x-2)^n
\end{aligned}
$$

Notice that we simplified the factorials in this case. You should always simplify them if there are more than one and it's possible to simplify them.

Also, do not get excited about the term sitting in front of the series. Sometimes we need to do that when we can't get a general formula that will hold for all values of *n*.

Example 8 Find the Taylor Series for $f\left(x\right)=\frac{1}{x^2}$ about $x=-1$.

Solution
Again, here are the derivatives and evaluations.

$$
\begin{aligned}
f^{(0)}\left(x\right) &= \frac{1}{x^2} & f^{(0)}\left(-1\right) &= \frac{1}{\left(-1\right)^2} = 1 \\
f^{(1)}\left(x\right) &= -\frac{2}{x^3} & f^{(1)}\left(-1\right) &= -\frac{2}{\left(-1\right)^3} = 2 \\
f^{(2)}\left(x\right) &= \frac{2\left(3\right)}{x^4} & f^{(2)}\left(-1\right) &= \frac{2\left(3\right)}{\left(-1\right)^4} = 2\left(3\right) \\
f^{(3)}\left(x\right) &= -\frac{2\left(3\right)\left(4\right)}{x^5} & f^{(3)}\left(-1\right) &= -\frac{2\left(3\right)\left(4\right)}{\left(-1\right)^5} = 2\left(3\right)\left(4\right) \\
&\vdots & &\vdots \\
f^{(n)}\left(x\right) &= \frac{\left(-1\right)^n\left(n+1\right)!}{x^{n+2}} & f^{(n)}\left(-1\right) &= \frac{\left(-1\right)^n\left(n+1\right)!}{\left(-1\right)^{n+2}} = \left(n+1\right)!
\end{aligned}
$$

Notice that all the negative signs will cancel out in the evaluation. Also, this formula will work for all *n*, unlike the previous example.

Here is the Taylor Series for this function.

$$
\begin{aligned}
\frac{1}{x^2} &= \sum_{n=0}^{\infty} \frac{f^{(n)}\left(-1\right)}{n!}\left(x+1\right)^n \\
&= \sum_{n=0}^{\infty} \frac{\left(n+1\right)!}{n!}\left(x+1\right)^n \\
&= \sum_{n=0}^{\infty} \left(n+1\right)\left(x+1\right)^n
\end{aligned}
$$

Now, let's work one of the easier examples in this section. The problem for most students is that it may not appear to be that easy (or maybe it will appear to be too easy) at first glance.

Example 9 Find the Taylor Series for $f\left(x\right)=x^3-10x^2+6$ about $x=3$.

Solution
Here are the derivatives for this problem.

$$
\begin{aligned}
f^{(0)}\left(x\right) &= x^3-10x^2+6 & f^{(0)}\left(3\right) &= -57 \\
f^{(1)}\left(x\right) &= 3x^2-20x & f^{(1)}\left(3\right) &= -33 \\
f^{(2)}\left(x\right) &= 6x-20 & f^{(2)}\left(3\right) &= -2 \\
f^{(3)}\left(x\right) &= 6 & f^{(3)}\left(3\right) &= 6 \\
f^{(n)}\left(x\right) &= 0 & f^{(4)}\left(3\right) &= 0 \qquad n \geq 4
\end{aligned}
$$

This Taylor series will terminate after $n=3$. This will always happen when we are finding the Taylor Series of a polynomial. Here is the Taylor Series for this one.

$$\begin{aligned} x^3-10x^2+6 &= \sum_{n=0}^{\infty} \frac{f^{(n)}(3)}{n!}(x-3)^n \\ &= f(3)+f'(3)(x-3)+\frac{f''(3)}{2!}(x-3)^2+\frac{f'''(3)}{3!}(x-3)^3+0 \\ &= -57-33(x-3)-(x-3)^2+(x-3)^3 \end{aligned}$$

When finding the Taylor Series of a polynomial we don't do any simplification of the right hand side. We leave it like it is. In fact, if we were to multiply everything out we just get back to the original polynomial!

While it's not apparent that writing the Taylor Series for a polynomial is useful there are times where this needs to be done. The problem is that they are beyond the scope of this course and so aren't covered here. For example, there is one application to series in the field of Differential Equations where this needs to be done on occasion.

So, we've seen quite a few examples of Taylor Series to this point and in all of them we were able to find general formulas for the series. This won't always be the case. To see an example of one that doesn't have a general formula check out the last example in the next section.

Before leaving this section there are three important Taylor Series that we've derived in this section that we should summarize up in one place. In my class I will assume that you know these formulas from this point on.

$$\mathbf{e}^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$$

Applications of Series

Now, that we know how to represent function as power series we can now talk about at least a couple of applications of series.

There are in fact many applications of series, unfortunately most of them are beyond the scope of this course. One application of power series (with the occasional use of Taylor Series) is in the field of Ordinary Differential Equations when finding Series Solutions to Differential Equations. If you are interested in seeing how that works you can check out that chapter of my Differential Equations notes.

Another application of series arises in the study of Partial Differential Equations. One of the more commonly used methods in that subject makes use of Fourier Series.

Many of the applications of series, especially those in the differential equations fields, rely on the fact that functions can be represented as a series. In these applications it is very difficult, if not impossible, to find the function itself. However, there are methods of determining the series representation for the unknown function.

While the differential equations applications are beyond the scope of this course there are some applications from a Calculus setting that we can look at.

Example 1 Determine a Taylor Series about $x = 0$ for the following integral.

$$\int \frac{\sin x}{x}\,dx$$

Solution

To do this we will first need to find a Taylor Series about $x = 0$ for the integrand. This however isn't terribly difficult. We already have a Taylor Series for sine about $x = 0$ so we'll just use that as follows,

$$\frac{\sin x}{x} = \frac{1}{x}\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n+1)!}$$

We can now do the problem.

$$\begin{aligned}\int \frac{\sin x}{x}\,dx &= \int \sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n+1)!}\,dx \\ &= C + \sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)(2n+1)!}\end{aligned}$$

So, while we can't integrate this function in terms of known functions we can come up with a series representation for the integral.

This idea of deriving a series representation for a function instead of trying to find the function itself is used quite often in several fields. In fact, there are some fields where this is one of the main ideas used and without this idea it would be very difficult to accomplish anything in those fields.

Another application of series isn't really an application of infinite series. It's more an application of partial sums. In fact, we've already seen this application in use once in this chapter. In the Estimating the Value of a Series we used a partial sum to estimate the value of a series. We can do the same thing with power series and series representations of functions. The main difference is that we will now be using the partial sum to approximate a function instead of a single value.

We will look at Taylor series for our examples, but we could just as easily use any series representation here. Recall that the **nth degree Taylor Polynomial of *f(x)*** is given by,

$$T_n(x) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i$$

Let's take a look at example of this.

Example 2 For the function $f(x) = \cos(x)$ plot the function as well as $T_2(x)$, $T_4(x)$, and $T_8(x)$ on the same graph for the interval [-4,4].

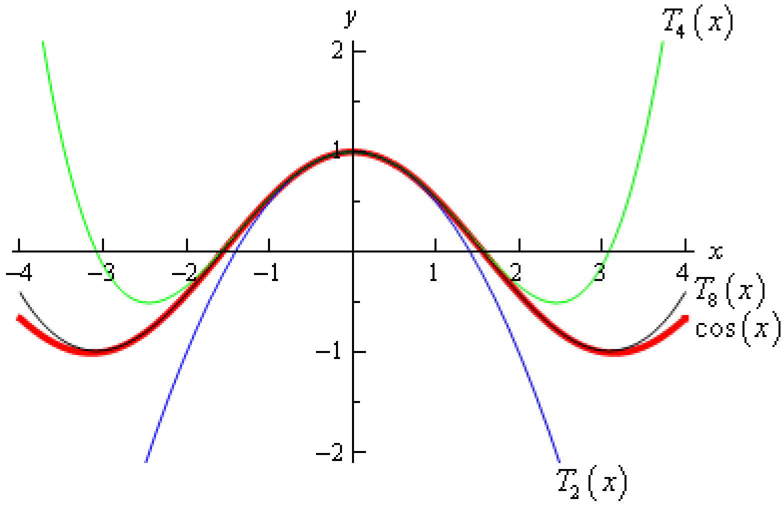
Solution
Here is the general formula for the Taylor polynomials for cosine.

$$T_{2n}(x) = \sum_{i=0}^{n} \frac{(-1)^i x^{2i}}{(2i)!}$$

The three Taylor polynomials that we've got are then,

$$T_2(x) = 1 - \frac{x^2}{2}$$
$$T_4(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24}$$
$$T_8(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + \frac{x^8}{40320}$$

Here is the graph of these three Taylor polynomials as well as the graph of cosine.



As we can see from this graph as we increase the degree of the Taylor polynomial it starts to look more and more like the function itself. In fact by the time we get to $T_8(x)$ the only difference is right at the ends. The higher the degree of the Taylor polynomial the better it approximates the function.

Also the larger the interval the higher degree Taylor polynomial we need to get a good approximation for the whole interval.

Before moving on let's write down a couple more Taylor polynomials from the previous example. Notice that because the Taylor series for cosine doesn't contain any terms with odd powers on *x* we get the following Taylor polynomials.

$$T_3(x) = 1 - \frac{x^2}{2}$$
$$T_5(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24}$$
$$T_9(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + \frac{x^8}{40320}$$

These are identical to those used in the example. Sometimes this will happen although that was not really the point of this. The point is to notice that the nth degree Taylor polynomial may actually have a degree that is less than *n*. It will never be more than *n*, but it can be less than *n*.

The final example in this section really isn't an application of series and probably belonged in the previous section. However, the previous section was getting too long so the example is in this section. This is an example of how to multiply series together and while this isn't an application of series it is something that does have to be done on occasion in the applications. So, in that sense it does belong in this section.

Example 3 Find the first three non-zero terms in the Taylor Series for $f(x) = \mathbf{e}^x \cos x$ about $x = 0$.

Solution
Before we start let's acknowledge that the easiest way to do this problem is to simply compute the first 3-4 derivatives, evaluate them at $x = 0$, plug into the formula and we'd be done. However, as we noted prior to this example we want to use this example to illustrate how we multiply series.

We will make use of the fact that we've got Taylor Series for each of these so we can use them in this problem.

$$\mathbf{e}^x \cos x = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}\right)$$

We're not going to completely multiply out these series. We're going to do enough of the multiplication to get an answer. The problem statement says that we want the first three *non-zero* terms. That non-zero bit is important as it is possible that some of the terms will be zero. If none of the terms are zero this would mean that the first three non-zero terms would be the constant term, x term, and x^2 term. However, because some might be zero let's assume that if we get all the terms up through x^4 we'll have enough to get the answer. If we've assumed wrong it will be very easy to fix so don't worry about that.

Now, let's write down the first few terms of each series and we'll stop at the x^4 term in each.

$$\mathbf{e}^x \cos x = \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots\right)\left(1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots\right)$$

Note that we do need to acknowledge that these series don't stop. That's the purpose of the "$+\cdots$" at the end of each. Just for a second however, let's suppose that each of these did stop and ask ourselves how we would multiply each out. If this were the case we would take every term in the second and multiply by every term in the first. In other words, we would first multiply every term in the second series by 1, then every term in the second series by x, then by x^2 *etc*.

By stopping each series at x^4 we have now guaranteed that we'll get all terms that have an exponent of 4 or less. Do you see why?

Each of the terms that we neglected to write down have an exponent of at least 5 and so multiplying by 1 or any power of x will result in a term with an exponent that is at a minimum 5. Therefore, none of the neglected terms will contribute terms with an exponent of 4 or less and so weren't needed.

So, let's start the multiplication process.

$$\mathbf{e}^x \cos x = \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots\right)\left(1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots\right)$$

$$= \underbrace{1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots}_{\text{Second Series} \times 1} + \underbrace{x - \frac{x^3}{2} + \frac{x^5}{24} + \cdots}_{\text{Second Series} \times x} + \underbrace{\frac{x^2}{2} - \frac{x^4}{4} + \frac{x^6}{48} + \cdots}_{\text{Second Series} \times x^2/2}$$

$$+ \underbrace{\frac{x^3}{6} - \frac{x^5}{12} + \frac{x^7}{144} + \cdots}_{\text{Second Series} \times x^3/6} + \underbrace{\frac{x^4}{24} - \frac{x^6}{48} + \frac{x^8}{576} + \cdots}_{\text{Second Series} \times x^4/24} + \cdots$$

Now, collect like terms ignoring everything with an exponent of 5 or more since we won't have all those terms and don't want them either. Doing this gives,

$$\mathbf{e}^x \cos x = 1 + x + \left(-\frac{1}{2} + \frac{1}{2}\right)x^2 + \left(-\frac{1}{2} + \frac{1}{6}\right)x^3 + \left(\frac{1}{24} - \frac{1}{4} + \frac{1}{24}\right)x^4 + \cdots$$

$$= 1 + x - \frac{x^3}{3} - \frac{x^4}{6} + \cdots$$

There we go. It looks like we over guessed and ended up with four non-zero terms, but that's okay. If we had under guessed and it turned out that we needed terms with x^5 in them all we would need to do at this point is go back and add in those terms to the original series and do a couple quick multiplications. In other words, there is no reason to completely redo all the work.

Binomial Series

In this final section of this chapter we are going to look at another series representation for a function. Before we do this let's first recall the following theorem.

Binomial Theorem

If n is any positive integer then,

$$\begin{aligned}\left(a+b\right)^n &= \sum_{i=0}^{n} \binom{n}{i} a^{n-i}\, b^i \\ &= a^n + na^{n-1}b + \frac{n\left(n-1\right)}{2!}a^{n-2}b^2 + \cdots + nab^{n-1} + b^n\end{aligned}$$

where,

$$\binom{n}{i} = \frac{n\left(n-1\right)\left(n-2\right)\cdots\left(n-i+1\right)}{i!} \qquad i = 1, 2, 3, \ldots n$$

$$\binom{n}{0} = 1$$

This is useful for expanding $\left(a+b\right)^n$ for large n when straight forward multiplication wouldn't be easy to do. Let's take a quick look at an example.

Example 1 Use the Binomial Theorem to expand $\left(2x-3\right)^4$

Solution

There really isn't much to do other than plugging into the theorem.

$$\begin{aligned}\left(2x-3\right)^4 &= \sum_{i=0}^{4} \binom{4}{i} \left(2x\right)^{4-i} \left(-3\right)^i \\ &= \binom{4}{0}\left(2x\right)^4 + \binom{4}{1}\left(2x\right)^3\left(-3\right) + \binom{4}{2}\left(2x\right)^2\left(-3\right)^2 + \binom{4}{3}\left(2x\right)\left(-3\right)^3 + \binom{4}{4}\left(-3\right)^4 \\ &= \left(2x\right)^4 + 4\left(2x\right)^3\left(-3\right) + \frac{4\left(3\right)}{2}\left(2x\right)^2\left(-3\right)^2 + 4\left(2x\right)\left(-3\right)^3 + \left(-3\right)^4 \\ &= 16x^4 - 96x^3 + 216x^2 - 216x + 81\end{aligned}$$

Now, the Binomial Theorem required that n be a positive integer. There is an extension to this however that allows for any number at all.

Binomial Series

If k is any number and $|x|<1$ then,

$$\left(1+x\right)^k = \sum_{n=0}^{\infty} \binom{k}{n} x^n$$
$$= 1 + kx + \frac{k\left(k-1\right)}{2!}x^2 + \frac{k\left(k-1\right)\left(k-2\right)}{3!}x^3 + \cdots$$

where,

$$\binom{k}{n} = \frac{k\left(k-1\right)\left(k-2\right)\cdots\left(k-n+1\right)}{n!} \qquad n = 1, 2, 3, \ldots$$
$$\binom{k}{0} = 1$$

So, similar to the binomial theorem except that it's an infinite series and we must have $|x|<1$ in order to get convergence.

Let's check out an example of this.

Example 2 Write down the first four terms in the binomial series for $\sqrt{9-x}$

Solution

So, in this case $k=\frac{1}{2}$ and we'll need to rewrite the term a little to put it into the form required.

$$\sqrt{9-x} = 3\left(1-\frac{x}{9}\right)^{\frac{1}{2}} = 3\left(1+\left(-\frac{x}{9}\right)\right)^{\frac{1}{2}}$$

The first four terms in the binomial series is then,

$$\sqrt{9-x} = 3\left(1+\left(-\frac{x}{9}\right)\right)^{\frac{1}{2}}$$
$$= 3\sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}\left(-\frac{x}{9}\right)^n$$
$$= 3\left[1+\left(\frac{1}{2}\right)\left(-\frac{x}{9}\right) + \frac{\frac{1}{2}\left(-\frac{1}{2}\right)}{2}\left(-\frac{x}{9}\right)^2 + \frac{\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{6}\left(-\frac{x}{9}\right)^3 + \cdots\right]$$
$$= 3 - \frac{x}{6} - \frac{x^2}{216} - \frac{x^3}{3888} - \cdots$$

Vectors

Introduction

This is a fairly short chapter. We will be taking a brief look at vectors and some of their properties. We will need some of this material in the next chapter and those of you heading on towards Calculus III will use a fair amount of this there as well.

Here is a list of topics in this chapter.

Vectors – The Basics

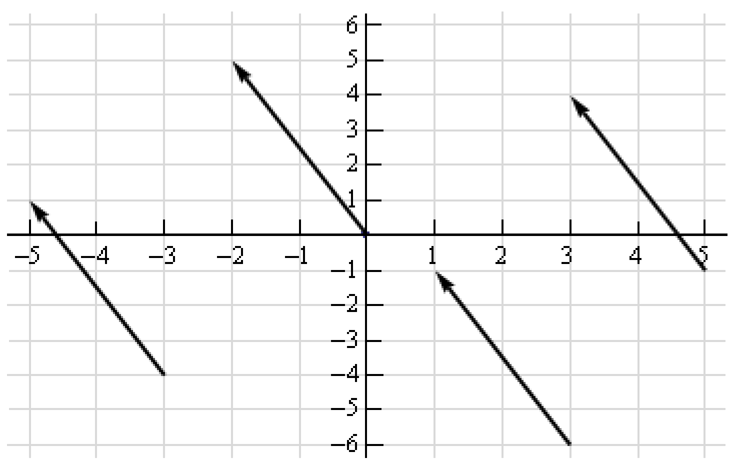
Let's start this section off with a quick discussion on what vectors are used for. Vectors are used to represent quantities that have both a magnitude and a direction. Good examples of quantities that can be represented by vectors are force and velocity. Both of these have a direction and a magnitude.

Let's consider force for a second. A force of say 5 Newtons that is applied in a particular direction can be applied at any point in space. In other words, the point where we apply the force does not change the force itself. Forces are independent of the point of application. To define a force all we need to know is the magnitude of the force and the direction that the force is applied in.

The same idea holds more generally with vectors. Vectors only impart magnitude and direction. They don't impart any information about where the quantity is applied. This is an important idea to always remember in the study of vectors.

In a graphical sense vectors are represented by directed line segments. The length of the line segment is the magnitude of the vector and the direction of the line segment is the direction of the vector. However, because vectors don't impart any information about where the quantity is applied any directed line segment with the same length and direction will represent the same vector.

Consider the sketch below.



Each of the directed line segments in the sketch represents the same vector. In each case the vector starts at a specific point then moves 2 units to the left and 5 units up. The notation that we'll use for this vector is,

$$\vec{v} = \left\langle -2, 5 \right\rangle$$

and each of the directed line segments in the sketch are called **representations** of the vector.

Be careful to distinguish vector notation, $\left\langle -2, 5 \right\rangle$, from the notation we use to represent coordinates of points, $\left(-2, 5 \right)$. The vector denotes a magnitude and a direction of a quantity while the point denotes a location in space. So don't mix the notations up!

A representation of the vector $\vec{v} = \langle a_1, a_2 \rangle$ in two dimensional space is any directed line segment, $\overrightarrow{AB}$, from the point $A = (x, y)$ to the point $B = (x + a_1, y + a_2)$. Likewise a representation of the vector $\vec{v} = \langle a_1, a_2, a_3 \rangle$ in three dimensional space is any directed line segment, $\overrightarrow{AB}$, from the point $A = (x, y, z)$ to the point $B = (x + a_1, y + a_2, z + a_3)$.

Note that there is very little difference between the two dimensional and three dimensional formulas above. To get from the three dimensional formula to the two dimensional formula all we did is take out the third component/coordinate. Because of this most of the formulas here are given only in their three dimensional version. If we need them in their two dimensional form we can easily modify the three dimensional form.

There is one representation of a vector that is special in some way. The representation of the vector $\vec{v} = \langle a_1, a_2, a_3 \rangle$ that starts at the point $A = (0,0,0)$ and ends at the point $B = (a_1, a_2, a_3)$ is called the **position vector** of the point (a_1, a_2, a_3). So, when we talk about position vectors we are specifying the initial and final point of the vector.

Position vectors are useful if we ever need to represent a point as a vector. As we'll see there are times in which we definitely are going to want to represent points as vectors. In fact, we're going to run into topics that can only be done if we represent points as vectors.

Next we need to discuss briefly how to generate a vector given the initial and final points of the representation. Given the two points $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$ the vector with the representation $\overrightarrow{AB}$ is,

$$\vec{v} = \langle b_1 - a_1, b_2 - a_2, b_3 - a_3 \rangle$$

Note that we have to be very careful with direction here. The vector above is the vector that starts at *A* and ends at *B*. The vector that starts at *B* and ends at *A*, *i.e.* with representation $\overrightarrow{BA}$ is,

$$\vec{w} = \langle a_1 - b_1, a_2 - b_2, a_3 - b_3 \rangle$$

These two vectors are different and so we do need to always pay attention to what point is the starting point and what point is the ending point. When determining the vector between two points we always subtract the initial point from the terminal point.

Example 1 Give the vector for each of the following.

(a) The vector from $(2, -7, 0)$ to $(1, -3, -5)$.

(b) The vector from $(1, -3, -5)$ to $(2, -7, 0)$.

(c) The position vector for $(-90, 4)$

Solution

(a) Remember that to construct this vector we subtract coordinates of the starting point from the ending point.

$$\langle 1-2, -3-(-7), -5-0 \rangle = \langle -1, 4, -5 \rangle$$

(b) Same thing here.

$$\langle 2-1, -7-(-3), 0-(-5) \rangle = \langle 1, -4, 5 \rangle$$

Notice that the only difference between the first two is the signs are all opposite. This difference is important as it is this difference that tells us that the two vectors point in opposite directions.

(c) Not much to this one other than acknowledging that the position vector of a point is nothing more than a vector with the point's coordinates as its components.

$$\langle -90, 4 \rangle$$

We now need to start discussing some of the basic concepts that we will run into on occasion.

Magnitude

The **magnitude**, or **length**, of the vector $\vec{v} = \langle a_1, a_2, a_3 \rangle$ is given by,

$$\|\vec{v}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

Example 2 Determine the magnitude of each of the following vectors.

(a) $\vec{a} = \langle 3, -5, 10 \rangle$

(b) $\vec{u} = \left\langle \frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}} \right\rangle$

(c) $\vec{w} = \langle 0, 0 \rangle$

(d) $\vec{i} = \langle 1, 0, 0 \rangle$

Solution

There isn't too much to these other than plug into the formula.

(a) $\|\vec{a}\| = \sqrt{9+25+100} = \sqrt{134}$

(b) $\|\vec{u}\| = \sqrt{\frac{1}{5}+\frac{4}{5}} = \sqrt{1} = 1$

(c) $\|\vec{w}\| = \sqrt{0+0} = 0$

(d) $\|\vec{i}\| = \sqrt{1+0+0} = 1$

We also have the following fact about the magnitude.

$$\text{If } \|\vec{a}\| = 0 \text{ then } \vec{a} = \vec{0}$$

This should make sense. Because we square all the components the only way we can get zero out of the formula was for the components to be zero in the first place.

Unit Vector

Any vector with magnitude of 1, *i.e.* $\|\vec{u}\| = 1$, is called a **unit vector**.

Example 3 Which of the vectors from Example 2 are unit vectors?

Solution
Both the second and fourth vectors had a length of 1 and so they are the only unit vectors from the first example.

Zero Vector
The vector $\vec{w} = \langle 0, 0 \rangle$ that we saw in the first example is called a zero vector since its components are all zero. Zero vectors are often denoted by $\vec{0}$. Be careful to distinguish 0 (the number) from $\vec{0}$ (the vector). The number 0 denotes the origin in space, while the vector $\vec{0}$ denotes a vector that has no magnitude or direction.

Standard Basis Vectors
The fourth vector from the second example, $\vec{i} = \langle 1, 0, 0 \rangle$, is called a **standard basis vector**. In three dimensional space there are three standard basis vectors,

$$\vec{i} = \langle 1, 0, 0 \rangle \qquad \vec{j} = \langle 0, 1, 0 \rangle \qquad \vec{k} = \langle 0, 0, 1 \rangle$$

In two dimensional space there are two standard basis vectors,

$$\vec{i} = \langle 1, 0 \rangle \qquad \vec{j} = \langle 0, 1 \rangle$$

Note that standard basis vectors are also unit vectors.

Warning
We are pretty much done with this section however, before proceeding to the next section we should point out that vectors are not restricted to two dimensional or three dimensional space. Vectors can exist in general n-dimensional space. The general notation for a n-dimensional vector is,

$$\vec{v} = \langle a_1, a_2, a_3, \ldots, a_n \rangle$$

and each of the a_i's are called **components** of the vector.

Because we will be working almost exclusively with two and three dimensional vectors in this course most of the formulas will be given for the two and/or three dimensional cases. However, most of the concepts/formulas will work with general vectors and the formulas are easily (and naturally) modified for general n-dimensional vectors. Also, because it is easier to visualize things in two dimensions most of the figures related to vectors will be two dimensional figures.

So, we need to be careful to not get too locked into the two or three dimensional cases from our discussions in this chapter. We will be working in these dimensions either because it's easier to visualize the situation or because physical restrictions of the problems will enforce a dimension upon us.

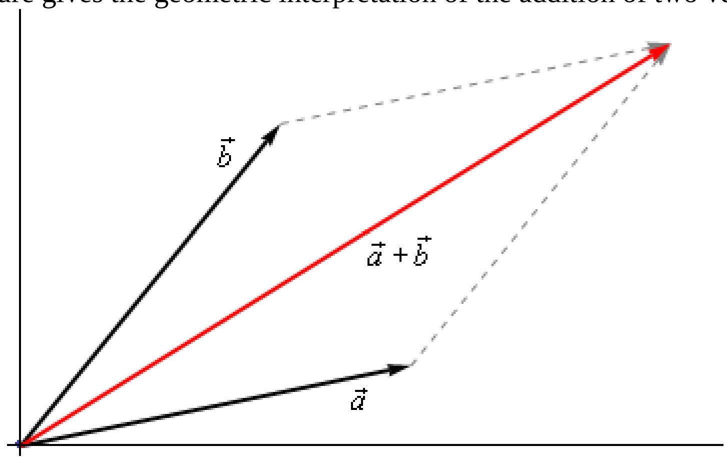
Vector Arithmetic

In this section we need to have a brief discussion of vector arithmetic.

We'll start with **addition** of two vectors. So, given the vectors $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\vec{b} = \langle b_1, b_2, b_3 \rangle$ the addition of the two vectors is given by the following formula.

$$\vec{a} + \vec{b} = \langle a_1 + b_1, a_2 + b_2, a_3 + b_3 \rangle$$

The following figure gives the geometric interpretation of the addition of two vectors.

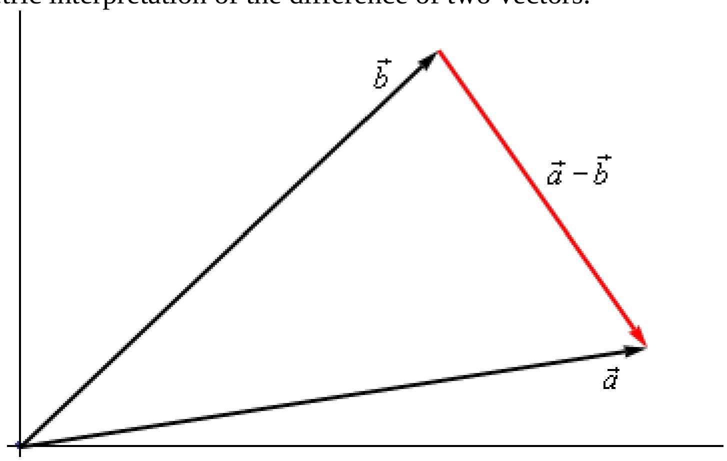


This is sometimes called the **parallelogram law** or **triangle law**.

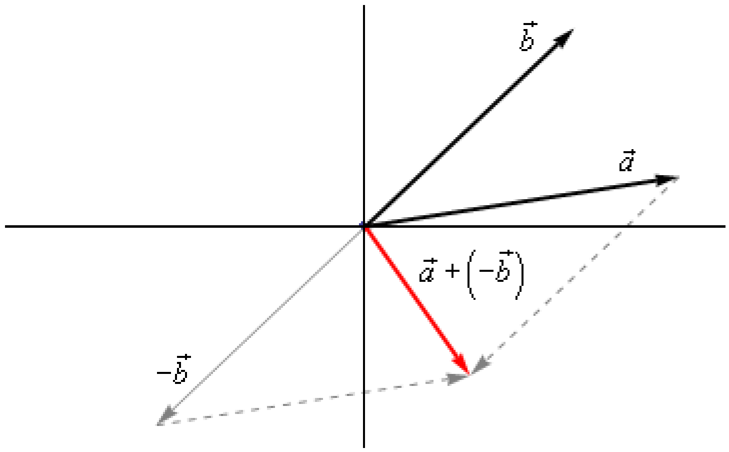
Computationally, **subtraction** is very similar. Given the vectors $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\vec{b} = \langle b_1, b_2, b_3 \rangle$ the difference of the two vectors is given by,

$$\vec{a} - \vec{b} = \langle a_1 - b_1, a_2 - b_2, a_3 - b_3 \rangle$$

Here is the geometric interpretation of the difference of two vectors.



It is a little harder to see this geometric interpretation. To help see this let's instead think of subtraction as the addition of $\vec{a}$ and $-\vec{b}$. First, as we'll see in a bit $-\vec{b}$ is the same vector as $\vec{b}$ with opposite signs on all the components. In other words, $-\vec{b}$ goes in the opposite direction as $\vec{b}$. Here is the vector set up for $\vec{a}+\left(-\vec{b}\right)$.



As we can see from this figure we can move the vector representing $\vec{a}+\left(-\vec{b}\right)$ to the position we've got in the first figure showing the difference of the two vectors.

Note that we can't add or subtract two vectors unless they have the same number of components. If they don't have the same number of components then addition and subtraction can't be done.

The next arithmetic operation that we want to look at is **scalar multiplication**. Given the vector $\vec{a}=\langle a_1,a_2,a_3\rangle$ and any number *c* the scalar multiplication is,

$$c\vec{a}=\langle ca_1,ca_2,ca_3\rangle$$

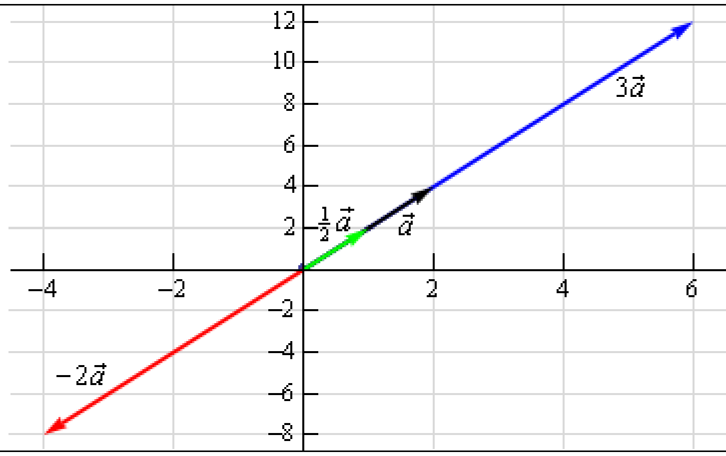
So, we multiply all the components by the constant *c*. To see the geometric interpretation of scalar multiplication let's take a look at an example.

Example 1 For the vector $\vec{a}=\langle 2,4\rangle$ compute $3\vec{a}$, $\frac{1}{2}\vec{a}$ and $-2\vec{a}$. Graph all four vectors on the same axis system.

Solution
Here are the three scalar multiplications.

$$3\vec{a}=\langle 6,12\rangle \qquad \frac{1}{2}\vec{a}=\langle 1,2\rangle \qquad -2\vec{a}=\langle -4,-8\rangle$$

Here is the graph for each of these vectors.



In the previous example we can see that if *c* is positive all scalar multiplication will do is stretch (if $c > 1$) or shrink (if $c < 1$) the original vector, but it won't change the direction. Likewise, if *c* is negative scalar multiplication will switch the direction so that the vector will point in exactly the opposite direction and it will again stretch or shrink the magnitude of the vector depending upon the size of *c*.

There are several nice applications of scalar multiplication that we should now take a look at.

The first is parallel vectors. This is a concept that we will see quite a bit over the next couple of sections. Two vectors are parallel if they have the same direction or are in exactly opposite directions. Now, recall again the geometric interpretation of scalar multiplication. When we performed scalar multiplication we generated new vectors that were parallel to the original vectors (and each other for that matter).

So, let's suppose that $\vec{a}$ and $\vec{b}$ are parallel vectors. If they are parallel then there must be a number *c* so that,

$$\vec{a} = c\vec{b}$$

So, two vectors are parallel if one is a scalar multiple of the other.

Example 2 Determine if the sets of vectors are parallel or not.

(a) $\vec{a} = \langle 2,-4,1\rangle$, $\vec{b} = \langle -6,12,-3\rangle$

(b) $\vec{a} = \langle 4,10\rangle$, $\vec{b} = \langle 2,-9\rangle$

Solution

(a) These two vectors are parallel since $\vec{b} = -3\vec{a}$

(b) These two vectors aren't parallel. This can be seen by noticing that $4\left(\frac{1}{2}\right) = 2$ and yet $10\left(\frac{1}{2}\right) = 5 \neq -9$. In other words we can't make $\vec{a}$ be a scalar multiple of $\vec{b}$.

The next application is best seen in an example.

Example 3 Find a unit vector that points in the same direction as $\vec{w} = \langle -5, 2, 1 \rangle$.

Solution
Okay, what we're asking for is a new parallel vector (points in the same direction) that happens to be a unit vector. We can do this with a scalar multiplication since all scalar multiplication does is change the length of the original vector (along with possibly flipping the direction to the opposite direction).

Here's what we'll do. First let's determine the magnitude of $\vec{w}$.

$$\|\vec{w}\| = \sqrt{25+4+1} = \sqrt{30}$$

Now, let's form the following new vector,

$$\vec{u} = \frac{1}{\|\vec{w}\|}\vec{w} = \frac{1}{\sqrt{30}}\langle -5, 2, 1 \rangle = \left\langle -\frac{5}{\sqrt{30}}, \frac{2}{\sqrt{30}}, \frac{1}{\sqrt{30}} \right\rangle$$

The claim is that this is a unit vector. That's easy enough to check

$$\|\vec{u}\| = \sqrt{\frac{25}{30}+\frac{4}{30}+\frac{1}{30}} = \sqrt{\frac{30}{30}} = 1$$

This vector also points in the same direction as $\vec{w}$ since it is only a scalar multiple of $\vec{w}$ and we used a positive multiple.

So, in general, given a vector $\vec{w}$, $\vec{u} = \dfrac{\vec{w}}{\|\vec{w}\|}$ will be a unit vector that points in the same direction as $\vec{w}$.

Standard Basis Vectors Revisited
In the previous section we introduced the idea of standard basis vectors without really discussing why they were important. We can now do that. Let's start with the vector

$$\vec{a} = \langle a_1, a_2, a_3 \rangle$$

We can use the addition of vectors to break this up as follows,

$$\begin{aligned}\vec{a} &= \langle a_1, a_2, a_3 \rangle \\ &= \langle a_1, 0, 0 \rangle + \langle 0, a_2, 0 \rangle + \langle 0, 0, a_3 \rangle\end{aligned}$$

Using scalar multiplication we can further rewrite the vector as,

$$\begin{aligned}\vec{a} &= \langle a_1, 0, 0 \rangle + \langle 0, a_2, 0 \rangle + \langle 0, 0, a_3 \rangle \\ &= a_1\langle 1, 0, 0 \rangle + a_2\langle 0, 1, 0 \rangle + a_3\langle 0, 0, 1 \rangle\end{aligned}$$

Finally, notice that these three new vectors are simply the three standard basis vectors for three dimensional space.

$$\langle a_1, a_2, a_3 \rangle = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$$

So, we can take any vector and write it in terms of the standard basis vectors. From this point on we will use the two notations interchangeably so make sure that you can deal with both notations.

Example 4 If $\vec{a} = \langle 3, -9, 1 \rangle$ and $\vec{w} = -\vec{i} + 8\vec{k}$ compute $2\vec{a} - 3\vec{w}$.

Solution
In order to do the problem we'll convert to one notation and then perform the indicated operations.

$$\begin{aligned} 2\vec{a} - 3\vec{w} &= 2\langle 3, -9, 1 \rangle - 3\langle -1, 0, 8 \rangle \\ &= \langle 6, -18, 2 \rangle - \langle -3, 0, 24 \rangle \\ &= \langle 9, -18, -22 \rangle \end{aligned}$$

We will leave this section with some basic properties of vector arithmetic.

Properties
If $\vec{v}$, $\vec{w}$ and $\vec{u}$ are vectors (each with the same number of components) and *a* and *b* are two numbers then we have the following properties.

$$\begin{aligned} &\vec{v} + \vec{w} = \vec{w} + \vec{v} && \vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w} \\ &\vec{v} + \vec{0} = \vec{v} && 1\vec{v} = \vec{v} \\ &a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w} && (a + b)\vec{v} = a\vec{v} + b\vec{v} \end{aligned}$$

The proofs of these are pretty much just "computation" proofs so we'll prove one of them and leave the others to you to prove.

Proof of $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$

We'll start with the two vectors, $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$ and $\vec{w} = \langle w_1, w_2, \ldots, w_n \rangle$ and yes we did mean for these to each have *n* components. The theorem works for general vectors so we may as well do the proof for general vectors.

Now, as noted above this is pretty much just a "computational" proof. What that means is that we'll compute the left side and then do some basic arithmetic on the result to show that we can make the left side look like the right side. Here is the work.

$$\begin{aligned} a(\vec{v} + \vec{w}) &= a\left(\langle v_1, v_2, \ldots, v_n \rangle + \langle w_1, w_2, \ldots, w_n \rangle\right) \\ &= a\langle v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n \rangle \\ &= \langle a(v_1 + w_1), a(v_2 + w_2), \ldots, a(v_n + w_n) \rangle \\ &= \langle av_1 + aw_1, av_2 + aw_2, \ldots, av_n + aw_n \rangle \\ &= \langle av_1, av_2, \ldots, av_n \rangle + \langle aw_1, aw_2, \ldots, aw_n \rangle \\ &= a\langle v_1, v_2, \ldots, v_n \rangle + a\langle w_1, w_2, \ldots, w_n \rangle = a\vec{v} + a\vec{w} \end{aligned}$$

■

Dot Product

The next topic for discussion is that of the dot product. Let's jump right into the definition of the dot product. Given the two vectors $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\vec{b} = \langle b_1, b_2, b_3 \rangle$ the dot product is,

$$\vec{a} \bullet \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \qquad (1)$$

Sometimes the dot product is called the **scalar product**. The dot product is also an example of an inner product and so on occasion you may hear it called an inner product.

Example 1 Compute the dot product for each of the following.

(a) $\vec{v} = 5\vec{i} - 8\vec{j}$, $\vec{w} = \vec{i} + 2\vec{j}$

(b) $\vec{a} = \langle 0, 3, -7 \rangle$, $\vec{b} = \langle 2, 3, 1 \rangle$

Solution

Not much to do with these other than use the formula.

(a) $\vec{v} \bullet \vec{w} = 5 - 16 = -11$

(b) $\vec{a} \bullet \vec{b} = 0 + 9 - 7 = 2$

Here are some properties of the dot product.

Properties

$$\vec{u} \bullet (\vec{v} + \vec{w}) = \vec{u} \bullet \vec{v} + \vec{u} \bullet \vec{w} \qquad (c\vec{v}) \bullet \vec{w} = \vec{v} \bullet (c\vec{w}) = c(\vec{v} \bullet \vec{w})$$

$$\vec{v} \bullet \vec{w} = \vec{w} \bullet \vec{v} \qquad \vec{v} \bullet \vec{0} = 0$$

$$\vec{v} \bullet \vec{v} = \|\vec{v}\|^2 \qquad \text{If } \vec{v} \bullet \vec{v} = 0 \text{ then } \vec{v} = \vec{0}$$

The proofs of these properties are mostly "computational" proofs and so we're only going to do a couple of them and leave the rest to you to prove.

Proof of $\vec{u} \bullet (\vec{v} + \vec{w}) = \vec{u} \bullet \vec{v} + \vec{u} \bullet \vec{w}$

We'll start with the three vectors, $\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$, $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$ and $\vec{w} = \langle w_1, w_2, \ldots, w_n \rangle$ and yes we did mean for these to each have *n* components. The theorem works for general vectors so we may as well do the proof for general vectors.

Now, as noted above this is pretty much just a "computational" proof. What that means is that we'll compute the left side and then do some basic arithmetic on the result to show that we can make the left side look like the right side. Here is the work.

$$
\begin{aligned}
\vec{u}\bullet(\vec{v}+\vec{w}) &= \langle u_1,u_2,\ldots,u_n\rangle\bullet(\langle v_1,v_2,\ldots,v_n\rangle+\langle w_1,w_2,\ldots,w_n\rangle) \\
&= \langle u_1,u_2,\ldots,u_n\rangle\bullet\langle v_1+w_1,v_2+w_2,\ldots,v_n+w_n\rangle \\
&= \langle u_1(v_1+w_1),u_2(v_2+w_2),\ldots,u_n(v_n+w_n)\rangle \\
&= \langle u_1v_1+u_1w_1,u_2v_2+u_2w_2,\ldots,u_nv_n+u_nw_n\rangle \\
&= \langle u_1v_1,u_2v_2,\ldots,u_nv_n\rangle+\langle u_1w_1,u_2w_2,\ldots,u_nw_n\rangle \\
&= \langle u_1,u_2,\ldots,u_n\rangle\bullet\langle v_1,v_2,\ldots,v_n\rangle+\langle u_1,u_2,\ldots,u_n\rangle\bullet\langle w_1,w_2,\ldots,w_n\rangle \\
&= \vec{u}\bullet\vec{v}+\vec{u}\bullet\vec{w}
\end{aligned}
$$

■

Proof of : If $\vec{v}\bullet\vec{v}=0$ **then** $\vec{v}=\vec{0}$

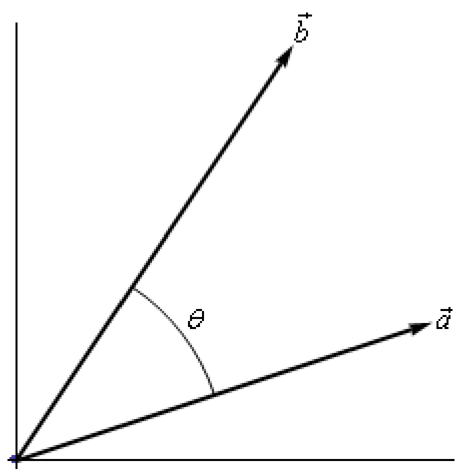
This is a pretty simple proof. Let’s start with $\vec{v}=\langle v_1,v_2,\ldots,v_n\rangle$ and compute the dot product.

$$
\begin{aligned}
\vec{v}\bullet\vec{v} &= \langle v_1,v_2,\ldots,v_n\rangle\bullet\langle v_1,v_2,\ldots,v_n\rangle \\
&= v_1^2+v_2^2+\cdots+v_n^2 \\
&= 0
\end{aligned}
$$

Now, since we know $v_i^2\geq 0$ for all *i* then the only way for this sum to be zero is to in fact have $v_i^2=0$. This in turn however means that we must have $v_i=0$ and so we must have had $\vec{v}=\vec{0}$.

■

There is also a nice geometric interpretation to the dot product. First suppose that θ is the angle between $\vec{a}$ and $\vec{b}$ such that $0\leq\theta\leq\pi$ as shown in the image below.



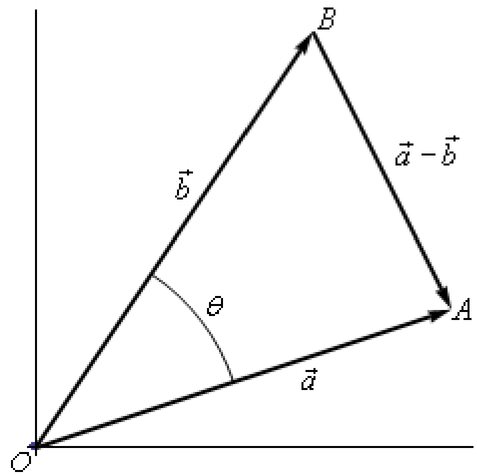
We can then have the following theorem.

Theorem

$$\vec{a}\bullet\vec{b} = \|\vec{a}\|\,\|\vec{b}\|\cos\theta \qquad (2)$$

Proof

Let's give a modified version of the sketch above.



The three vectors above form the triangle *AOB* and note that the length of each side is nothing more than the magnitude of the vector forming that side.

The Law of Cosines tells us that,

$$\|\vec{a}-\vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\,\|\vec{b}\|\cos\theta$$

Also using the properties of dot products we can write the left side as,

$$\begin{aligned}\|\vec{a}-\vec{b}\|^2 &= \left(\vec{a}-\vec{b}\right)\bullet\left(\vec{a}-\vec{b}\right)\\ &= \vec{a}\bullet\vec{a} - \vec{a}\bullet\vec{b} - \vec{b}\bullet\vec{a} + \vec{b}\bullet\vec{b}\\ &= \|\vec{a}\|^2 - 2\vec{a}\bullet\vec{b} + \|\vec{b}\|^2\end{aligned}$$

Our original equation is then,

$$\begin{aligned}\|\vec{a}-\vec{b}\|^2 &= \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\,\|\vec{b}\|\cos\theta\\ \|\vec{a}\|^2 - 2\vec{a}\bullet\vec{b} + \|\vec{b}\|^2 &= \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\,\|\vec{b}\|\cos\theta\\ -2\vec{a}\bullet\vec{b} &= -2\|\vec{a}\|\,\|\vec{b}\|\cos\theta\\ \vec{a}\bullet\vec{b} &= \|\vec{a}\|\,\|\vec{b}\|\cos\theta\end{aligned}$$

■

The formula from this theorem is often used not to compute a dot product but instead to find the angle between two vectors. Note as well that while the sketch of the two vectors in the proof is for two dimensional vectors the theorem is valid for vectors of any dimension (as long as they have the same dimension of course).

Let's see an example of this.

Example 2 Determine the angle between $\vec{a} = \langle 3, -4, -1 \rangle$ and $\vec{b} = \langle 0, 5, 2 \rangle$.

Solution

We will need the dot product as well as the magnitudes of each vector.

$$\vec{a} \bullet \vec{b} = -22 \qquad \|\vec{a}\| = \sqrt{26} \qquad \|\vec{b}\| = \sqrt{29}$$

The angle is then,

$$\cos\theta = \frac{\vec{a} \bullet \vec{b}}{\|\vec{a}\|\,\|\vec{b}\|} = \frac{-22}{\sqrt{26}\sqrt{29}} = -0.8011927$$

$$\theta = \cos^{-1}\left(-0.8011927\right) = 2.5 \text{ radians} = 143.24 \text{ degrees}$$

The dot product gives us a very nice method for determining if two vectors are perpendicular and it will give another method for determining when two vectors are parallel. Note as well that often we will use the term **orthogonal** in place of perpendicular.

Now, if two vectors are orthogonal then we know that the angle between them is 90 degrees. From (2) this tells us that if two vectors are orthogonal then,

$$\vec{a} \bullet \vec{b} = 0$$

Likewise, if two vectors are parallel then the angle between them is either 0 degrees (pointing in the same direction) or 180 degrees (pointing in the opposite direction). Once again using (2) this would mean that one of the following would have to be true.

$$\vec{a} \bullet \vec{b} = \|\vec{a}\|\,\|\vec{b}\| \; (\theta = 0^\circ) \qquad \text{OR} \qquad \vec{a} \bullet \vec{b} = -\|\vec{a}\|\,\|\vec{b}\| \; (\theta = 180^\circ)$$

Example 3 Determine if the following vectors are parallel, orthogonal, or neither.

(a) $\vec{a} = \langle 6, -2, -1 \rangle$, $\vec{b} = \langle 2, 5, 2 \rangle$

(b) $\vec{u} = 2\vec{i} - \vec{j}$, $\vec{v} = -\frac{1}{2}\vec{i} + \frac{1}{4}\vec{j}$

Solution

(a) First get the dot product to see if they are orthogonal.

$$\vec{a} \bullet \vec{b} = 12 - 10 - 2 = 0$$

The two vectors are orthogonal.

(b) Again, let's get the dot product first.

$$\vec{u}\bullet\vec{v} = -1 - \frac{1}{4} = -\frac{5}{4}$$

So, they aren't orthogonal. Let's get the magnitudes and see if they are parallel.

$$\|\vec{u}\| = \sqrt{5} \qquad \|\vec{v}\| = \sqrt{\frac{5}{16}} = \frac{\sqrt{5}}{4}$$

Now, notice that,

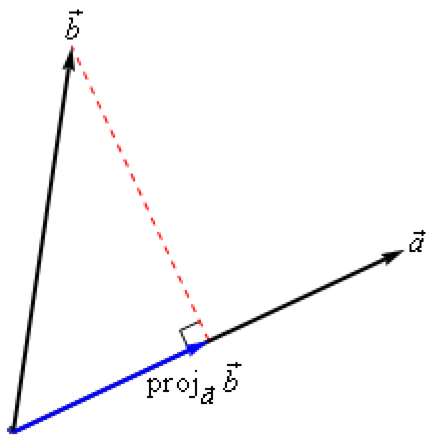
$$\vec{u}\bullet\vec{v} = -\frac{5}{4} = -\sqrt{5}\left(\frac{\sqrt{5}}{4}\right) = -\|\vec{u}\|\,\|\vec{v}\|$$

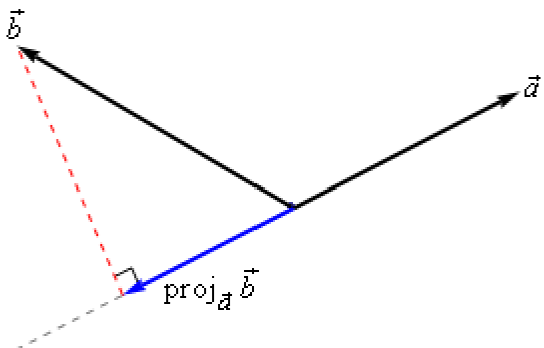
So, the two vectors are parallel.

There are several nice applications of the dot product as well that we should look at.

Projections

The best way to understand projections is to see a couple of sketches. So, given two vectors $\vec{a}$ and $\vec{b}$ we want to determine the projection of $\vec{b}$ onto $\vec{a}$. The projection is denoted by $\text{proj}_{\vec{a}}\,\vec{b}$. Here are a couple of sketches illustrating the projection.





So, to get the projection of $\vec{b}$ onto $\vec{a}$ we drop straight down from the end of $\vec{b}$ until we hit (and form a right angle) with the line that is parallel to $\vec{a}$. The projection is then the vector that is parallel to $\vec{a}$, starts at the same point both of the original vectors started at and ends where the dashed line hits the line parallel to $\vec{a}$.

There is an nice formula for finding the projection of $\vec{b}$ onto $\vec{a}$. Here it is,

$$\text{proj}_{\vec{a}}\,\vec{b} = \frac{\vec{a}\bullet\vec{b}}{\|\vec{a}\|^2}\,\vec{a}$$

Note that we also need to be very careful with notation here. The projection of $\vec{a}$ onto $\vec{b}$ is given by

$$\text{proj}_{\vec{b}}\,\vec{a} = \frac{\vec{a}\bullet\vec{b}}{\left\|\vec{b}\right\|^2}\vec{b}$$

We can see that this will be a totally different vector. This vector is parallel to $\vec{b}$, while $\text{proj}_{\vec{a}}\,\vec{b}$ is parallel to $\vec{a}$. So, be careful with notation and make sure you are finding the correct projection.

Here's an example.

Example 4 Determine the projection of $\vec{b} = \langle 2,1,-1\rangle$ onto $\vec{a} = \langle 1,0,-2\rangle$.

Solution
We need the dot product and the magnitude of $\vec{a}$.

$$\vec{a}\bullet\vec{b} = 4 \qquad\qquad \left\|\vec{a}\right\|^2 = 5$$

The projection is then,

$$\begin{aligned}\text{proj}_{\vec{a}}\,\vec{b} &= \frac{\vec{a}\bullet\vec{b}}{\left\|\vec{a}\right\|^2}\vec{a}\\ &= \frac{4}{5}\langle 1,0,-2\rangle\\ &= \left\langle \frac{4}{5},0,-\frac{8}{5}\right\rangle\end{aligned}$$

For comparison purposes let's do it the other way around as well.

Example 5 Determine the projection of $\vec{a} = \langle 1,0,-2\rangle$ onto $\vec{b} = \langle 2,1,-1\rangle$.

Solution
We need the dot product and the magnitude of $\vec{b}$.

$$\vec{a}\bullet\vec{b} = 4 \qquad\qquad \left\|\vec{b}\right\|^2 = 6$$

The projection is then,

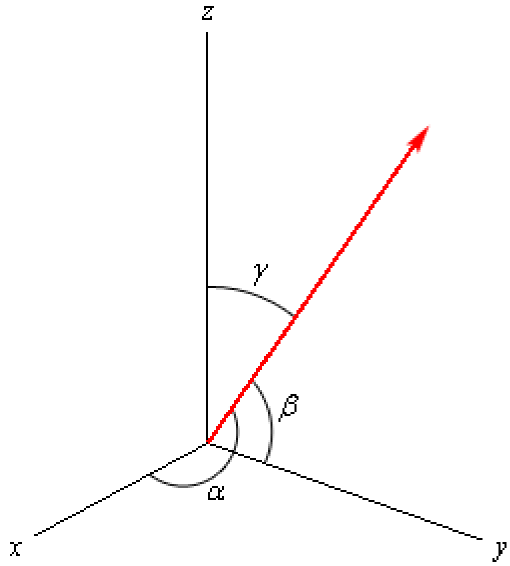
$$\begin{aligned}\text{proj}_{\vec{b}}\,\vec{a} &= \frac{\vec{a}\bullet\vec{b}}{\left\|\vec{b}\right\|^2}\vec{b}\\ &= \frac{4}{6}\langle 2,1,-1\rangle\\ &= \left\langle \frac{4}{3},\frac{2}{3},-\frac{2}{3}\right\rangle\end{aligned}$$

As we can see from the previous two examples the two projections are different so be careful.

Direction Cosines
This application of the dot product requires that we be in three dimensional space unlike all the other applications we've looked at to this point.

Let's start with a vector, $\vec{a}$, in three dimensional space. This vector will form angles with the *x*-axis (α), the *y*-axis (β), and the *z*-axis (γ). These angles are called **direction angles** and the cosines of these angles are called **direction cosines**.

Here is a sketch of a vector and the direction angles.



The formulas for the direction cosines are,

$$\cos\alpha = \frac{\vec{a}\bullet\vec{i}}{\|\vec{a}\|} = \frac{a_1}{\|\vec{a}\|} \qquad \cos\beta = \frac{\vec{a}\bullet\vec{j}}{\|\vec{a}\|} = \frac{a_2}{\|\vec{a}\|} \qquad \cos\gamma = \frac{\vec{a}\bullet\vec{k}}{\|\vec{a}\|} = \frac{a_3}{\|\vec{a}\|}$$

where $\vec{i}$, $\vec{j}$ and $\vec{k}$ are the standard basis vectors.

Let's verify the first dot product above. We'll leave the rest to you to verify.

$$\vec{a}\bullet\vec{i} = \langle a_1, a_2, a_3\rangle \bullet \langle 1, 0, 0\rangle = a_1$$

Here are a couple of nice facts about the direction cosines.

1. The vector $\vec{u} = \langle \cos\alpha, \cos\beta, \cos\gamma\rangle$ is a unit vector.
2. $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$
3. $\vec{a} = \|\vec{a}\|\langle \cos\alpha, \cos\beta, \cos\gamma\rangle$

Let's do a quick example involving direction cosines.

Example 6 Determine the direction cosines and direction angles for $\vec{a} = \left\langle 2, 1, -4 \right\rangle$.

Solution
We will need the magnitude of the vector.

$$\left\| \vec{a} \right\| = \sqrt{4+1+16} = \sqrt{21}$$

The direction cosines and angles are then,

$$\cos\alpha = \frac{2}{\sqrt{21}} \qquad \alpha = 1.119 \text{ radians} = 64.123 \text{ degrees}$$

$$\cos\beta = \frac{1}{\sqrt{21}} \qquad \beta = 1.351 \text{ radians} = 77.396 \text{ degrees}$$

$$\cos\gamma = \frac{-4}{\sqrt{21}} \qquad \gamma = 2.632 \text{ radians} = 150.794 \text{ degrees}$$

Cross Product

In this final section of this chapter we will look at the cross product of two vectors. We should note that the cross product requires both of the vectors to be three dimensional vectors.

Also, before getting into how to compute these we should point out a major difference between dot products and cross products. The result of a dot product is a number and the result of a cross product is a vector! Be careful not to confuse the two.

So, let's start with the two vectors $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\vec{b} = \langle b_1, b_2, b_3 \rangle$ then the cross product is given by the formula,

$$\vec{a} \times \vec{b} = \langle a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1 \rangle$$

This is not an easy formula to remember. There are two ways to derive this formula. Both of them use the fact that the cross product is really the determinant of a 3x3 matrix. If you don't know what this is that is don't worry about it. You don't need to know anything about matrices or determinants to use either of the methods. The notation for the determinant is as follows,

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

The first row is the standard basis vectors and must appear in the order given here. The second row is the components of $\vec{a}$ and the third row is the components of $\vec{b}$. Now, let's take a look at the different methods for getting the formula.

The first method uses the Method of Cofactors. If you don't know the method of cofactors that is fine, the result is all that we need. Here is the formula.

$$\vec{a} \times \vec{b} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \vec{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \vec{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \vec{k}$$

where,

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

This formula is not as difficult to remember as it might at first appear to be. First, the terms alternate in sign and notice that the 2x2 is missing the column below the standard basis vector that multiplies it as well as the row of standard basis vectors.

The second method is slightly easier; however, many textbooks don't cover this method as it will only work on 3x3 determinants. This method says to take the determinant as listed above and then copy the first two columns onto the end as shown below.

$$\vec{a}\times\vec{b}=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{vmatrix}\begin{matrix}\vec{i} & \vec{j}\\ a_1 & a_2\\ b_1 & b_2\end{matrix}$$

We now have three diagonals that move from left to right and three diagonals that move from right to left. We multiply along each diagonal and add those that move from left to right and subtract those that move from right to left.

This is best seen in an example. We'll also use this example to illustrate a fact about cross products.

Example 1 If $\vec{a}=\langle 2,1,-1\rangle$ and $\vec{b}=\langle -3,4,1\rangle$ compute each of the following.

(a) $\vec{a}\times\vec{b}$

(b) $\vec{b}\times\vec{a}$

Solution

(a) Here is the computation for this one.

$$\begin{aligned}\vec{a}\times\vec{b}&=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ 2 & 1 & -1\\ -3 & 4 & 1\end{vmatrix}\begin{matrix}\vec{i} & \vec{j}\\ 2 & 1\\ -3 & 4\end{matrix}\\ &=\vec{i}\,(1)(1)+\vec{j}\,(-1)(-3)+\vec{k}\,(2)(4)-\vec{j}\,(2)(1)-\vec{i}\,(-1)(4)-\vec{k}\,(1)(-3)\\ &=5\vec{i}+\vec{j}+11\vec{k}\end{aligned}$$

(b) And here is the computation for this one.

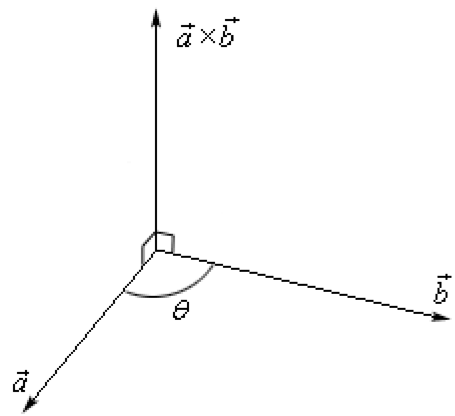
$$\begin{aligned}\vec{b}\times\vec{a}&=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ -3 & 4 & 1\\ 2 & 1 & -1\end{vmatrix}\begin{matrix}\vec{i} & \vec{j}\\ -3 & 4\\ 2 & 1\end{matrix}\\ &=\vec{i}\,(4)(-1)+\vec{j}\,(1)(2)+\vec{k}\,(-3)(1)-\vec{j}\,(-3)(-1)-\vec{i}\,(1)(1)-\vec{k}\,(4)(2)\\ &=-5\vec{i}-\vec{j}-11\vec{k}\end{aligned}$$

Notice that switching the order of the vectors in the cross product simply changed all the signs in the result. Note as well that this means that the two cross products will point in exactly opposite directions since they only differ by a sign. We'll formalize up this fact shortly when we list several facts.

There is also a geometric interpretation of the cross product. First we will let θ be the angle between the two vectors $\vec{a}$ and $\vec{b}$ and assume that $0\le\theta\le\pi$, then we have the following fact,

$$\left\|\vec{a}\times\vec{b}\right\|=\left\|\vec{a}\right\|\left\|\vec{b}\right\|\sin\theta \qquad (1)$$

and the following figure.



There should be a natural question at this point. How did we know that the cross product pointed in the direction that we've given it here?

First, as this figure, implies the cross product is orthogonal to both of the original vectors. This will always be the case with one exception that we'll get to in a second.

Second, we knew that it pointed in the upward direction (in this case) by the "right hand rule". This says that if we take our right hand, start at $\vec{a}$ and rotate our fingers towards $\vec{b}$ our thumb will point in the direction of the cross product. Therefore, if we'd sketched in $\vec{b} \times \vec{a}$ above we would have gotten a vector in the downward direction.

Example 2 A plane is defined by any three points that are in the plane. If a plane contains the points $P = (1,0,0)$, $Q = (1,1,1)$ and $R = (2,-1,3)$ find a vector that is orthogonal to the plane.

Solution

The one way that we know to get an orthogonal vector is to take a cross product. So, if we could find two vectors that we knew were in the plane and took the cross product of these two vectors we know that the cross product would be orthogonal to both the vectors. However, since both the vectors are in the plane the cross product would then also be orthogonal to the plane.

So, we need two vectors that are in the plane. This is where the points come into the problem. Since all three points lie in the plane any vector between them must also be in the plane. There are many ways to get two vectors between these points. We will use the following two,

$$\overrightarrow{PQ} = \langle 1-1, 1-0, 1-0 \rangle = \langle 0,1,1 \rangle$$

$$\overrightarrow{PR} = \langle 2-1, -1-0, 3-0 \rangle = \langle 1,-1,3 \rangle$$

The cross product of these two vectors will be orthogonal to the plane. So, let's find the cross product.

$$\overrightarrow{PQ} \times \overrightarrow{PR} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & 1 & 1 \\ 1 & -1 & 3 \end{vmatrix} \begin{matrix} \vec{i} & \vec{j} \\ 0 & 1 \\ 1 & -1 \end{matrix}$$

$$= 4\vec{i} + \vec{j} - \vec{k}$$

So, the vector $4\vec{i} + \vec{j} - \vec{k}$ will be orthogonal to the plane containing the three points.

Now, let's address the one time where the cross product will not be orthogonal to the original vectors. If the two vectors, $\vec{a}$ and $\vec{b}$, are parallel then the angle between them is either 0 or 180 degrees. From (1) this implies that,

$$\left\| \vec{a} \times \vec{b} \right\| = 0$$

From a fact about the magnitude we saw in the first section we know that this implies

$$\vec{a} \times \vec{b} = \vec{0}$$

In other words, it won't be orthogonal to the original vectors since we have the zero vector. This does give us another test for parallel vectors however.

Fact

If $\vec{a} \times \vec{b} = \vec{0}$ then $\vec{a}$ and $\vec{b}$ will be parallel vectors.

Let's also formalize up the fact about the cross product being orthogonal to the original vectors.

Fact

Provided $\vec{a} \times \vec{b} \neq \vec{0}$ then $\vec{a} \times \vec{b}$ is orthogonal to both $\vec{a}$ and $\vec{b}$.

Here are some nice properties about the cross product.

Properties

If $\vec{u}$, $\vec{v}$ and $\vec{w}$ are vectors and c is a number then,

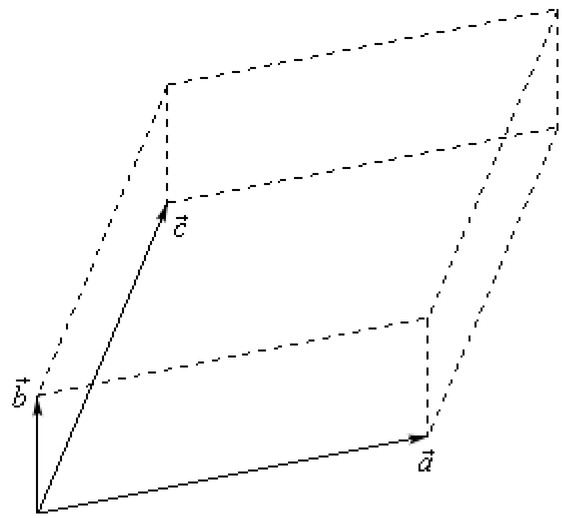
$$\vec{u} \times \vec{v} = -\vec{v} \times \vec{u} \qquad \left(c\vec{u}\right) \times \vec{v} = \vec{u} \times \left(c\vec{v}\right) = c\left(\vec{u} \times \vec{v}\right)$$

$$\vec{u} \times \left(\vec{v} + \vec{w}\right) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w} \qquad \vec{u} \bullet \left(\vec{v} \times \vec{w}\right) = \left(\vec{u} \times \vec{v}\right) \bullet \vec{w}$$

$$\vec{u} \bullet \left(\vec{v} \times \vec{w}\right) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

The determinant in the last fact is computed in the same way that the cross product is computed. We will see an example of this computation shortly.

There are a couple of geometric applications to the cross product as well. Suppose we have three vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ and we form the three dimensional figure shown below.



The area of the parallelogram (two dimensional front of this object) is given by,

$$\text{Area} = \left\| \vec{a} \times \vec{b} \right\|$$

and the volume of the parallelepiped (the whole three dimensional object) is given by,

$$\text{Volume} = \left| \vec{a} \bullet \left(\vec{b} \times \vec{c} \right) \right|$$

Note that the absolute value bars are required since the quantity could be negative and volume isn't negative.

We can use this volume fact to determine if three vectors lie in the same plane or not. If three vectors lie in the same plane then the volume of the parallelepiped will be zero.

Example 3 Determine if the three vectors $\vec{a} = \langle 1, 4, -7 \rangle$, $\vec{b} = \langle 2, -1, 4 \rangle$ and $\vec{c} = \langle 0, -9, 18 \rangle$ lie in the same plane or not.

Solution
So, as we noted prior to this example all we need to do is compute the volume of the parallelepiped formed by these three vectors. If the volume is zero they lie in the same plane and if the volume isn't zero they don't lie in the same plane.

$$\vec{a}\bullet\left(\vec{b}\times\vec{c}\right)=\begin{vmatrix}1 & 4 & -7\\ 2 & -1 & 4\\ 0 & -9 & 18\end{vmatrix}\begin{matrix}1 & 4\\ 2 & -1\\ 0 & -9\end{matrix}$$

$$\begin{aligned}&=(1)(-1)(18)+(4)(4)(0)+(-7)(2)(-9)-\\&\qquad(4)(2)(18)-(1)(4)(-9)-(-7)(-1)(0)\\&=-18+126-144+36\\&=0\end{aligned}$$

So, the volume is zero and so they lie in the same plane.

Three Dimensional Space

Introduction

In this chapter we will start taking a more detailed look at three dimensional space (3-D space or $\mathbb{R}^3$). This is a very important topic for Calculus III since a good portion of Calculus III is done in three (or higher) dimensional space.

We will be looking at the equations of graphs in 3-D space as well as vector valued functions and how we do calculus with them. We will also be taking a look at a couple of new coordinate systems for 3-D space.

This is the only chapter that exists in two places in my notes. When I originally wrote these notes all of these topics were covered in Calculus II however, we have since moved several of them into Calculus III. So, rather than split the chapter up I have kept it in the Calculus II notes and also put a copy in the Calculus III notes. Many of the sections not covered in Calculus III will be used on occasion there anyway and so they serve as a quick reference for when we need them.

Here is a list of topics in this chapter.

The 3-D Coordinate System – We will introduce the concepts and notation for the three dimensional coordinate system in this section.

Equations of Lines – In this section we will develop the various forms for the equation of lines in three dimensional space.

Equations of Planes – Here we will develop the equation of a plane.

Quadric Surfaces – In this section we will be looking at some examples of quadric surfaces.

Functions of Several Variables – A quick review of some important topics about functions of several variables.

Vector Functions – We introduce the concept of vector functions in this section. We concentrate primarily on curves in three dimensional space. We will however, touch briefly on surfaces as well.

Calculus with Vector Functions – Here we will take a quick look at limits, derivatives, and integrals with vector functions.

Tangent, Normal and Binormal Vectors – We will define the tangent, normal and binormal vectors in this section.

Arc Length with Vector Functions – In this section we will find the arc length of a vector function.

Curvature – We will determine the curvature of a function in this section.

Velocity and Acceleration – In this section we will revisit a standard application of derivatives. We will look at the velocity and acceleration of an object whose position function is given by a vector function.

Cylindrical Coordinates – We will define the cylindrical coordinate system in this section. The cylindrical coordinate system is an alternate coordinate system for the three dimensional coordinate system.

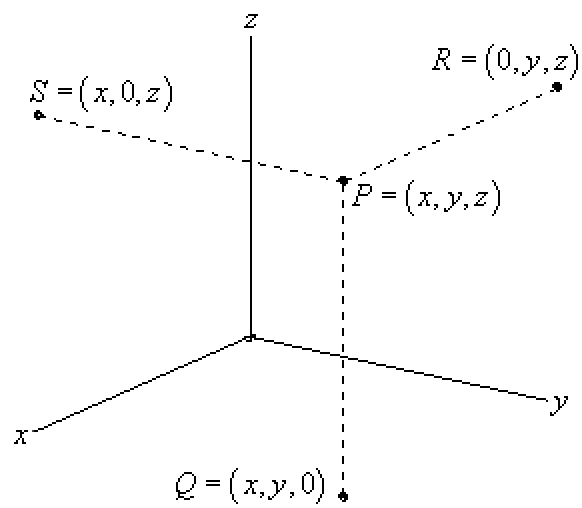
Spherical Coordinates – In this section we will define the spherical coordinate system. The spherical coordinate system is yet another alternate coordinate system for the three dimensional coordinate system.

The 3-D Coordinate System

We'll start the chapter off with a fairly short discussion introducing the 3-D coordinate system and the conventions that we'll be using. We will also take a brief look at how the different coordinate systems can change the graph of an equation.

Let's first get some basic notation out of the way. The 3-D coordinate system is often denoted by $\mathbb{R}^3$. Likewise the 2-D coordinate system is often denoted by $\mathbb{R}^2$ and the 1-D coordinate system is denoted by $\mathbb{R}$. Also, as you might have guessed then a general *n* dimensional coordinate system is often denoted by $\mathbb{R}^n$.

Next, let's take a quick look at the basic coordinate system.



This is the standard placement of the axes in this class. It is assumed that only the positive directions are shown by the axes. If we need the negative axes for any reason we will put them in as needed.

Also note the various points on this sketch. The point *P* is the general point sitting out in 3-D space. If we start at *P* and drop straight down until we reach a *z*-coordinate of zero we arrive at the point *Q*. We say that *Q* sits in the *xy*-plane. The *xy*-plane corresponds to all the points which have a zero *z*-coordinate. We can also start at *P* and move in the other two directions as shown to get points in the *xz*-plane (this is *S* with a *y*-coordinate of zero) and the *yz*-plane (this is *R* with an *x*-coordinate of zero).

Collectively, the *xy*, *xz*, and *yz*-planes are sometimes called the coordinate planes. In the remainder of this class you will need to be able to deal with the various coordinate planes so make sure that you can.

Also, the point *Q* is often referred to as the projection of *P* in the *xy*-plane. Likewise, *R* is the projection of *P* in the *yz*-plane and *S* is the projection of *P* in the *xz*-plane.

Many of the formulas that you are used to working with in $\mathbb{R}^2$ have natural extensions in $\mathbb{R}^3$. For instance the distance between two points in $\mathbb{R}^2$ is given by,

$$d\left(P_1,P_2\right)=\sqrt{\left(x_2-x_1\right)^2+\left(y_2-y_1\right)^2}$$

While the distance between any two points in $\mathbb{R}^3$ is given by,

$$d\left(P_1,P_2\right)=\sqrt{\left(x_2-x_1\right)^2+\left(y_2-y_1\right)^2+\left(z_2-z_1\right)^2}$$

Likewise, the general equation for a circle with center $\left(h,k\right)$ and radius *r* is given by,

$$\left(x-h\right)^2+\left(y-k\right)^2=r^2$$

and the general equation for a sphere with center $\left(h,k,l\right)$ and radius *r* is given by,

$$\left(x-h\right)^2+\left(y-k\right)^2+\left(z-l\right)^2=r^2$$

With that said we do need to be careful about just translating everything we know about $\mathbb{R}^2$ into $\mathbb{R}^3$ and assuming that it will work the same way. A good example of this is in graphing to some extent. Consider the following example.

Example 1 Graph $x=3$ in $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^3$.

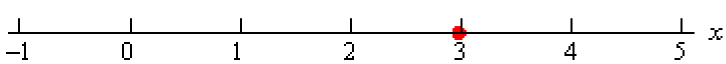
Solution

In $\mathbb{R}$ we have a single coordinate system and so $x=3$ is a point in a 1-D coordinate system.

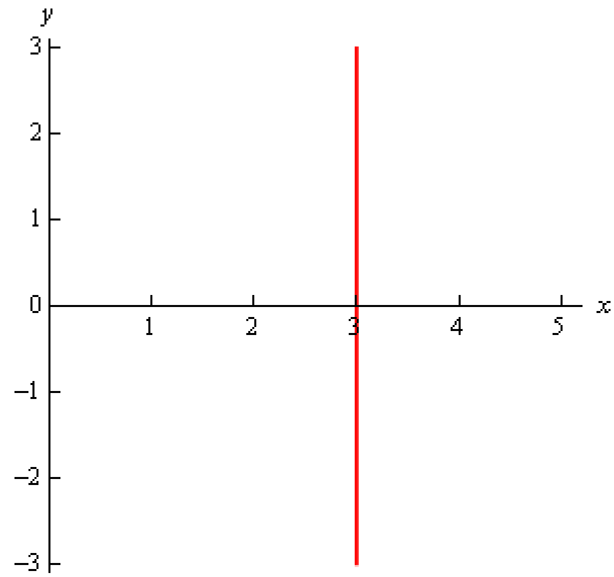
In $\mathbb{R}^2$ the equation $x=3$ tells us to graph all the points that are in the form $\left(3,y\right)$. This is a vertical line in a 2-D coordinate system.

In $\mathbb{R}^3$ the equation $x=3$ tells us to graph all the points that are in the form $\left(3,y,z\right)$. If you go back and look at the coordinate plane points this is very similar to the coordinates for the *yz*-plane except this time we have $x=3$ instead of $x=0$. So, in a 3-D coordinate system this is a plane that will be parallel to the *yz*-plane and pass through the *x*-axis at $x=3$.

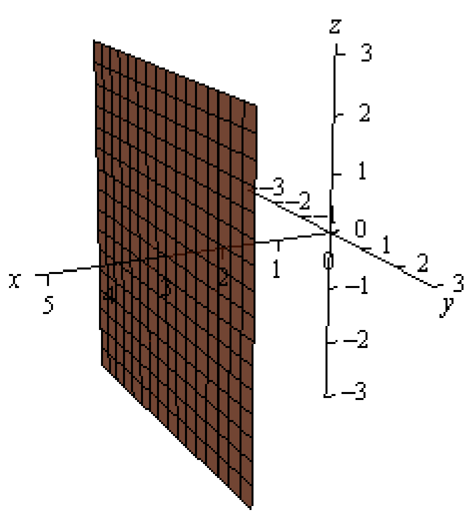
Here is the graph of $x=3$ in $\mathbb{R}$.

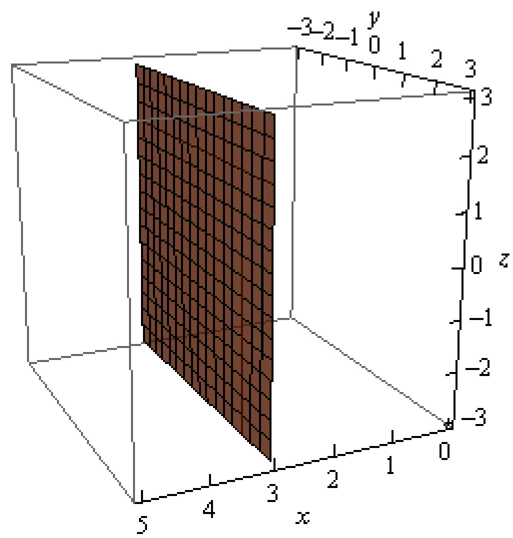


Here is the graph of $x=3$ in $\mathbb{R}^2$.



Finally, here is the graph of $x = 3$ in $\mathbb{R}^3$. Note that we've presented this graph in two different styles. On the left we've got the traditional axis system that we're used to seeing and on the right we've put the graph in a box. Both views can be convenient on occasion to help with perspective and so we'll often do this with 3D graphs and sketches.





Note that at this point we can now write down the equations for each of the coordinate planes as well using this idea.

$$
\begin{array}{ll}
z = 0 & xy - \text{plane} \\
y = 0 & xz - \text{plane} \\
x = 0 & yz - \text{plane}
\end{array}
$$

Let's take a look at a slightly more general example.

Example 2 Graph $y = 2x - 3$ in $\mathbb{R}^2$ and $\mathbb{R}^3$.

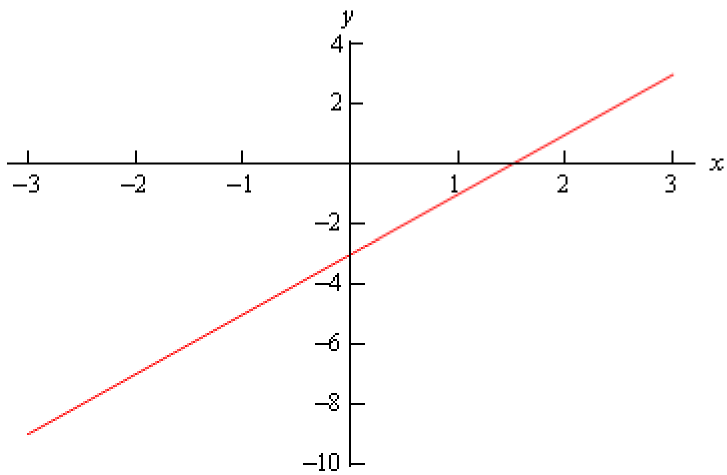
Solution

Of course we had to throw out $\mathbb{R}$ for this example since there are two variables which means that we can't be in a 1-D space.

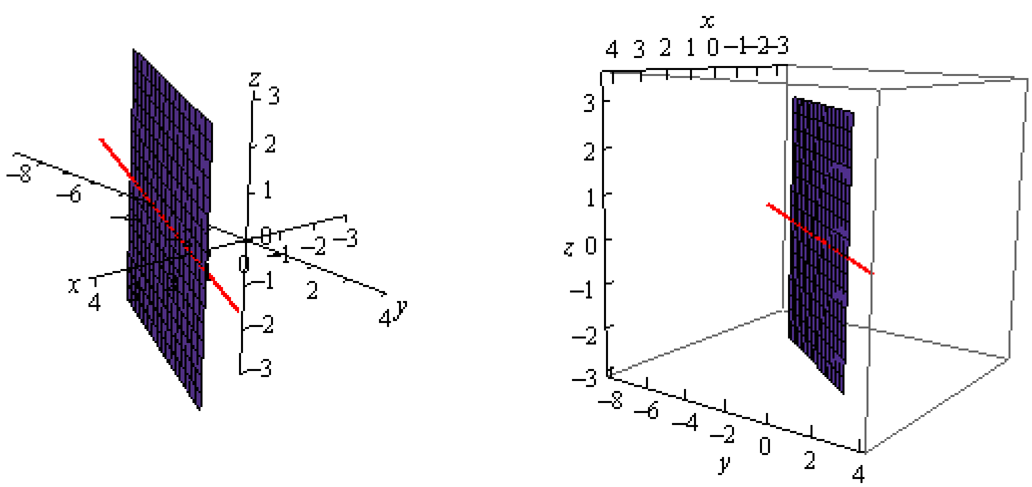
In $\mathbb{R}^2$ this is a line with slope 2 and a *y* intercept of -3.

However, in $\mathbb{R}^3$ this is not necessarily a line. Because we have not specified a value of *z* we are forced to let *z* take any value. This means that at any particular value of *z* we will get a copy of this line. So, the graph is then a vertical plane that lies over the line given by $y = 2x - 3$ in the *xy*-plane.

Here is the graph in $\mathbb{R}^2$.



here is the graph in $\mathbb{R}^3$.



Notice that if we look to where the plane intersects the *xy*-plane we will get the graph of the line in $\mathbb{R}^2$ as noted in the above graph by the red line through the plane.

Let's take a look at one more example of the difference between graphs in the different coordinate systems.

Example 3 Graph $x^2 + y^2 = 4$ in $\mathbb{R}^2$ and $\mathbb{R}^3$.

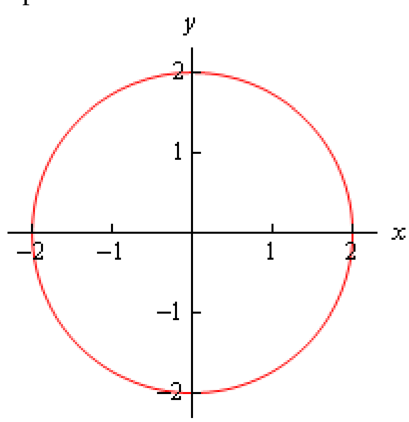
Solution

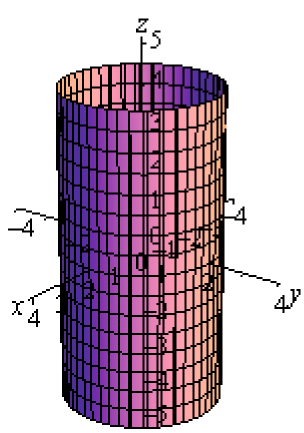
As with the previous example this won't have a 1-D graph since there are two variables.

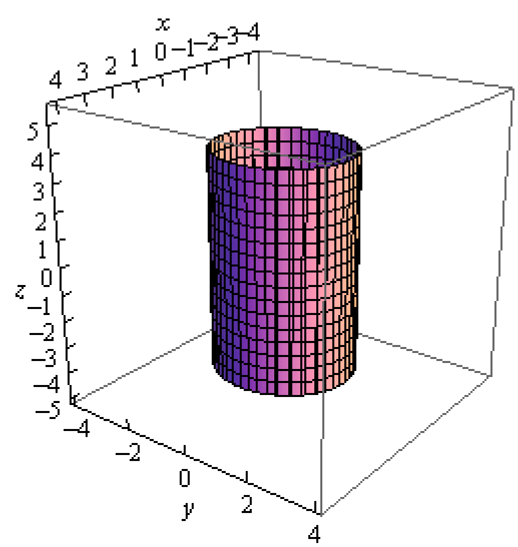
In $\mathbb{R}^2$ this is a circle centered at the origin with radius 2.

In $\mathbb{R}^3$ however, as with the previous example, this may or may not be a circle. Since we have not specified *z* in any way we must assume that *z* can take on any value. In other words, at any value of *z* this equation must be satisfied and so at any value *z* we have a circle of radius 2 centered on the *z*-axis. This means that we have a cylinder of radius 2 centered on the *z*-axis.

Here are the graphs for this example.







Notice that again, if we look to where the cylinder intersects the *xy*-plane we will again get the circle from $\mathbb{R}^2$.

We need to be careful with the last two examples. It would be tempting to take the results of these and say that we can't graph lines or circles in $\mathbb{R}^3$ and yet that doesn't really make sense. There is no reason for there to not be graphs of lines or circles in $\mathbb{R}^3$. Let's think about the example of the circle. To graph a circle in $\mathbb{R}^3$ we would need to do something like $x^2 + y^2 = 4$ at $z = 5$. This would be a circle of radius 2 centered on the *z*-axis at the level of $z = 5$. So, as long as we specify a *z* we will get a circle and not a cylinder. We will see an easier way to specify circles in a later section.

We could do the same thing with the line from the second example. However, we will be looking at lines in more generality in the next section and so we'll see a better way to deal with lines in $\mathbb{R}^3$ there.

The point of the examples in this section is to make sure that we are being careful with graphing equations and making sure that we always remember which coordinate system that we are in.

Another quick point to make here is that, as we've seen in the above examples, many graphs of equations in $\mathbb{R}^3$ are surfaces. That doesn't mean that we can't graph curves in $\mathbb{R}^3$. We can and will graph curves in $\mathbb{R}^3$ as well as we'll see later in this chapter.

Equations of Lines

In this section we need to take a look at the equation of a line in $\mathbb{R}^3$. As we saw in the previous section the equation $y = mx + b$ does not describe a line in $\mathbb{R}^3$, instead it describes a plane. This doesn't mean however that we can't write down an equation for a line in 3-D space. We're just going to need a new way of writing down the equation of a curve.

So, before we get into the equations of lines we first need to briefly look at vector functions. We're going to take a more in depth look at vector functions later. At this point all that we need to worry about is notational issues and how they can be used to give the equation of a curve.

The best way to get an idea of what a vector function is and what its graph looks like is to look at an example. So, consider the following vector function.

$$\vec{r}(t) = \langle t, 1 \rangle$$

A vector function is a function that takes one or more variables, one in this case, and returns a vector. Note as well that a vector function can be a function of two or more variables. However, in those cases the graph may no longer be a curve in space.

The vector that the function gives can be a vector in whatever dimension we need it to be. In the example above it returns a vector in $\mathbb{R}^2$. When we get to the real subject of this section, equations of lines, we'll be using a vector function that returns a vector in $\mathbb{R}^3$

Now, we want to determine the graph of the vector function above. In order to find the graph of our function we'll think of the vector that the vector function returns as a position vector for points on the graph. Recall that a position vector, say $\vec{v} = \langle a, b \rangle$, is a vector that starts at the origin and ends at the point (a, b).

So, to get the graph of a vector function all we need to do is plug in some values of the variable and then plot the point that corresponds to each position vector we get out of the function and play connect the dots. Here are some evaluations for our example.

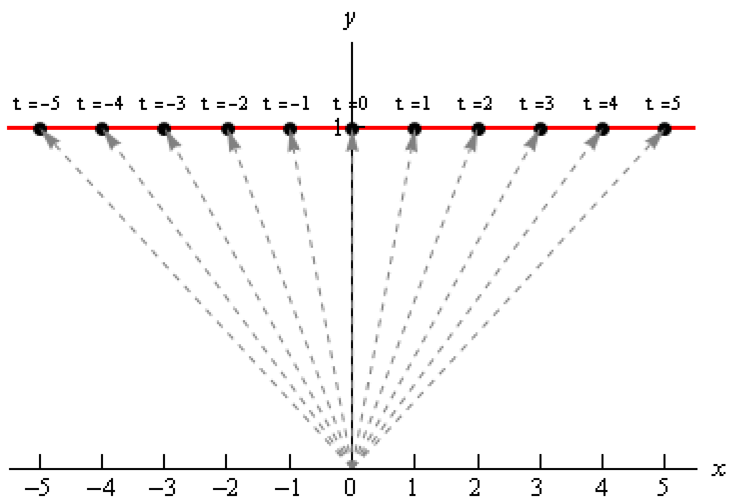
$$\vec{r}(-3) = \langle -3, 1 \rangle \qquad \vec{r}(-1) = \langle -1, 1 \rangle \qquad \vec{r}(2) = \langle 2, 1 \rangle \qquad \vec{r}(5) = \langle 5, 1 \rangle$$

So, each of these are position vectors representing points on the graph of our vector function. The points,

$$(-3,1) \qquad (-1,1) \qquad (2,1) \qquad (5,1)$$

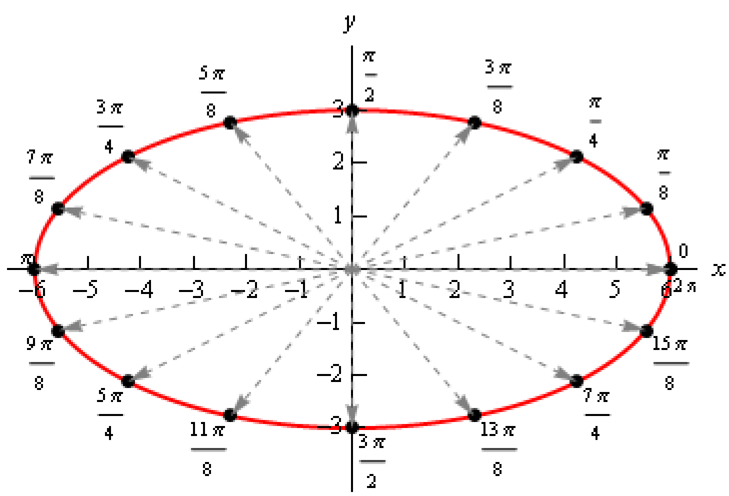
are all points that lie on the graph of our vector function.

If we do some more evaluations and plot all the points we get the following sketch.



In this sketch we've included the position vector (in gray and dashed) for several evaluations as well as the *t* (above each point) we used for each evaluation. It looks like, in this case the graph of the vector equation is in fact the line $y = 1$.

Here's another quick example. Here is the graph of $\vec{r}(t) = \langle 6\cos t, 3\sin t \rangle$.



In this case we get an ellipse. It is important to not come away from this section with the idea that vector functions only graph out lines. We'll be looking at lines in this section, but the graphs of vector functions do not have to be lines as the example above shows.

We'll leave this brief discussion of vector functions with another way to think of the graph of a vector function. Imagine that a pencil/pen is attached to the end of the position vector and as we increase the variable the resulting position vector moves and as it moves the pencil/pen on the end sketches out the curve for the vector function.

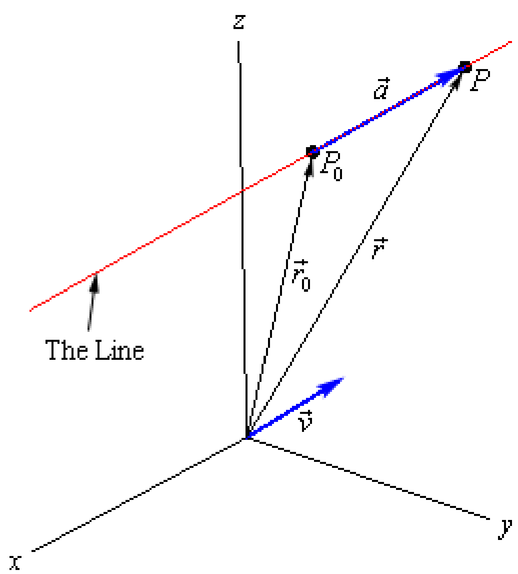
Okay, we now need to move into the actual topic of this section. We want to write down the equation of a line in $\mathbb{R}^3$ and as suggested by the work above we will need a vector function to do this. To see how we're going to do this let's think about what we need to write down the equation of a line in $\mathbb{R}^2$. In two dimensions we need the slope (*m*) and a point that was on the line in order to write down the equation.

In $\mathbb{R}^3$ that is still all that we need except in this case the "slope" won't be a simple number as it was in two dimensions. In this case we will need to acknowledge that a line can have a three dimensional slope. So, we need something that will allow us to describe a direction that is potentially in three dimensions. We already have a quantity that will do this for us. Vectors give directions and can be three dimensional objects.

So, let's start with the following information. Suppose that we know a point that is on the line, $P_0 = \left(x_0, y_0, z_0\right)$, and that $\vec{v} = \langle a, b, c \rangle$ is some vector that is parallel to the line. Note, in all likelihood, $\vec{v}$ will not be on the line itself. We only need $\vec{v}$ to be parallel to the line. Finally, let $P = \left(x, y, z\right)$ be any point on the line.

Now, since our "slope" is a vector let's also represent the two points on the line as vectors. We'll do this with position vectors. So, let $\vec{r_0}$ and $\vec{r}$ be the position vectors for P_0 and P respectively. Also, for no apparent reason, let's define $\vec{a}$ to be the vector with representation $\overrightarrow{P_0 P}$.

We now have the following sketch with all these points and vectors on it.



Now, we've shown the parallel vector, $\vec{v}$, as a position vector but it doesn't need to be a position vector. It can be anywhere, a position vector, on the line or off the line, it just needs to be parallel to the line.

Next, notice that we can write $\vec{r}$ as follows,

$$\vec{r} = \vec{r_0} + \vec{a}$$

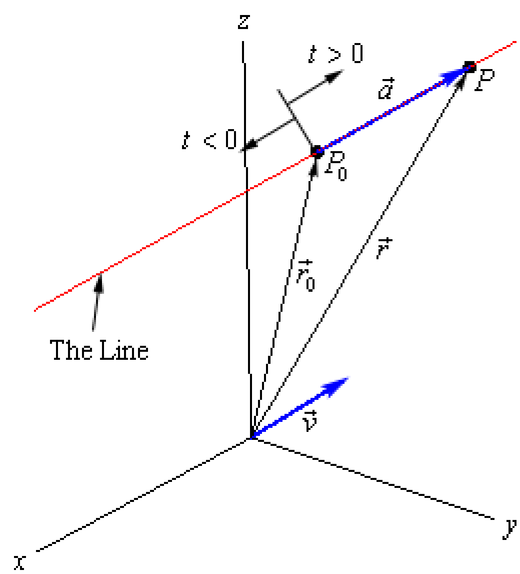
If you're not sure about this go back and check out the sketch for vector addition in the [vector arithmetic] section. Now, notice that the vectors $\vec{a}$ and $\vec{v}$ are parallel. [Therefore] there is a number, *t*, such that

$$\vec{a} = t\,\vec{v}$$

We now have,

$$\vec{r} = \overrightarrow{r_0} + t\,\vec{v} = \langle x_0, y_0, z_0 \rangle + t\langle a, b, c \rangle$$

This is called the **vector form of the equation of a line**. The only part of this equation that is not known is the t. Notice that $t\,\vec{v}$ will be a vector that lies along the line and it tells us how far from the original point that we should move. If t is positive we move away from the original point in the direction of $\vec{v}$ (right in our sketch) and if t is negative we move away from the original point in the opposite direction of $\vec{v}$ (left in our sketch). As t varies over all possible values we will completely cover the line. The following sketch shows this dependence on t of our sketch.



There are several other forms of the equation of a line. To get the first alternate form let's start with the vector form and do a slight rewrite.

$$\begin{aligned} \vec{r} &= \langle x_0, y_0, z_0 \rangle + t\langle a, b, c \rangle \\ \langle x, y, z \rangle &= \langle x_0 + ta, y_0 + tb, z_0 + tc \rangle \end{aligned}$$

The only way for two vectors to be equal is for the components to be equal. In other words,

$$\begin{aligned} x &= x_0 + ta \\ y &= y_0 + tb \\ z &= z_0 + tc \end{aligned}$$

This set of equations is called the **parametric form of the equation of a line**. Notice as well that this is really nothing more than an extension of the parametric equations we've seen previously. The only difference is that we are now working in three dimensions instead of two dimensions.

To get a point on the line all we do is pick a t and plug into either form of the line. In the vector form of the line we get a position vector for the point and in the parametric form we get the actual coordinates of the point.

There is one more form of the line that we want to look at. If we assume that a, b, and c are all non-zero numbers we can solve each of the equations in the parametric form of the line for t. We can then set all of them equal to each other since t will be the same number in each. Doing this gives the following,

$$\frac{x-x_0}{a}=\frac{y-y_0}{b}=\frac{z-z_0}{c}$$

This is called the **symmetric equations of the line**.

If one of a, b, or c does happen to be zero we can still write down the symmetric equations. To see this let's suppose that $b=0$. In this case t will not exist in the parametric equation for y and so we will only solve the parametric equations for x and z for t. We then set those equal and acknowledge the parametric equation for y as follows,

$$\frac{x-x_0}{a}=\frac{z-z_0}{c} \qquad\qquad y=y_0$$

Let's take a look at an example.

Example 1 Write down the equation of the line that passes through the points $\left(2,-1,3\right)$ and $\left(1,4,-3\right)$. Write down all three forms of the equation of the line.

Solution

To do this we need the vector $\vec{v}$ that will be parallel to the line. This can be any vector as long as it's parallel to the line. In general, $\vec{v}$ won't lie on the line itself. However, in this case it will. All we need to do is let $\vec{v}$ be the vector that starts at the second point and ends at the first point. Since these two points are on the line the vector between them will also lie on the line and will hence be parallel to the line. So,

$$\vec{v}=\left\langle 1,-5,6\right\rangle$$

Note that the order of the points was chosen to reduce the number of minus signs in the vector. We could just have easily gone the other way.

Once we've got $\vec{v}$ there really isn't anything else to do. To use the vector form we'll need a point on the line. We've got two and so we can use either one. We'll use the first point. Here is the vector form of the line.

$$\vec{r}=\left\langle 2,-1,3\right\rangle+t\left\langle 1,-5,6\right\rangle=\left\langle 2+t,-1-5t,3+6t\right\rangle$$

Once we have this equation the other two forms follow. Here are the parametric equations of the line.

$$x = 2 + t$$
$$y = -1 - 5t$$
$$z = 3 + 6t$$

Here is the symmetric form.

$$\frac{x-2}{1} = \frac{y+1}{-5} = \frac{z-3}{6}$$

Example 2 Determine if the line that passes through the point $(0,-3,8)$ and is parallel to the line given by $x = 10 + 3t$, $y = 12t$ and $z = -3 - t$ passes through the *xz*-plane. If it does give the coordinates of that point.

Solution

To answer this we will first need to write down the equation of the line. We know a point on the line and just need a parallel vector. We know that the new line must be parallel to the line given by the parametric equations in the problem statement. That means that any vector that is parallel to the given line must also be parallel to the new line.

Now recall that in the parametric form of the line the numbers multiplied by *t* are the components of the vector that is parallel to the line. Therefore, the vector,

$$\vec{v} = \langle 3, 12, -1 \rangle$$

is parallel to the given line and so must also be parallel to the new line.

The equation of new line is then,

$$\vec{r} = \langle 0,-3,8 \rangle + t\langle 3,12,-1 \rangle = \langle 3t, -3+12t, 8-t \rangle$$

If this line passes through the *xz*-plane then we know that the *y*-coordinate of that point must be zero. So, let's set the *y* component of the equation equal to zero and see if we can solve for *t*. If we can, this will give the value of *t* for which the point will pass through the *xz*-plane.

$$-3 + 12t = 0 \qquad \Rightarrow \qquad t = \frac{1}{4}$$

So, the line does pass through the *xz*-plane. To get the complete coordinates of the point all we need to do is plug $t = \frac{1}{4}$ into any of the equations. We'll use the vector form.

$$\vec{r} = \left\langle 3\left(\frac{1}{4}\right), -3 + 12\left(\frac{1}{4}\right), 8 - \frac{1}{4} \right\rangle = \left\langle \frac{3}{4}, 0, \frac{31}{4} \right\rangle$$

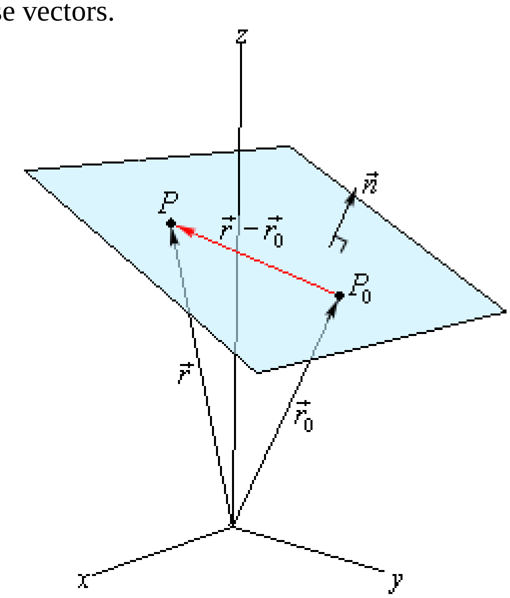
Recall that this vector is the position vector for the point on the line and so the coordinates of the point where the line will pass through the *xz*-plane are $\left(\frac{3}{4}, 0, \frac{31}{4}\right)$.

Equations of Planes

In the first section of this chapter we saw a couple of equations of planes. However, none of those equations had three variables in them and were really extensions of graphs that we could look at in two dimensions. We would like a more general equation for planes.

So, let's start by assuming that we know a point that is on the plane, $P_0 = (x_0, y_0, z_0)$. Let's also suppose that we have a vector that is orthogonal (perpendicular) to the plane, $\vec{n} = \langle a, b, c \rangle$. This vector is called the **normal vector**. Now, assume that $P = (x, y, z)$ is any point in the plane. Finally, since we are going to be working with vectors initially we'll let $\vec{r_0}$ and $\vec{r}$ be the position vectors for P_0 and P respectively.

Here is a sketch of all these vectors.



Notice that we added in the vector $\vec{r} - \vec{r_0}$ which will lie completely in the plane. Also notice that we put the normal vector on the plane, but there is actually no reason to expect this to be the case. We put it here to illustrate the point. It is completely possible that the normal vector does not touch the plane in any way.

Now, because $\vec{n}$ is orthogonal to the plane, it's also orthogonal to any vector that lies in the plane. In particular it's orthogonal to $\vec{r} - \vec{r_0}$. Recall from the Dot Product section that two orthogonal vectors will have a dot product of zero. In other words,

$$\vec{n} \bullet \left(\vec{r} - \vec{r_0} \right) = 0 \qquad \Rightarrow \qquad \vec{n} \bullet \vec{r} = \vec{n} \bullet \vec{r_0}$$

This is called the **vector equation of the plane**.

A slightly more useful form of the equations is as follows. Start with the first form of the vector equation and write down a vector for the difference.

$$\langle a,b,c\rangle \bullet \left(\langle x,y,z\rangle - \langle x_0,y_0,z_0\rangle\right) = 0$$
$$\langle a,b,c\rangle \bullet \langle x-x_0, y-y_0, z-z_0\rangle = 0$$

Now, actually compute the dot product to get,

$$a\left(x-x_0\right)+b\left(y-y_0\right)+c\left(z-z_0\right)=0$$

This is called the **scalar equation of plane**. Often this will be written as,

$$ax+by+cz=d$$

where $d = ax_0 + by_0 + cz_0$.

This second form is often how we are given equations of planes. Notice that if we are given the equation of a plane in this form we can quickly get a normal vector for the plane. A normal vector is,

$$\vec{n} = \langle a,b,c\rangle$$

Let's work a couple of examples.

Example 1 Determine the equation of the plane that contains the points $P=\left(1,-2,0\right)$, $Q=\left(3,1,4\right)$ and $R=\left(0,-1,2\right)$.

Solution

In order to write down the equation of plane we need a point (we've got three so we're cool there) and a normal vector. We need to find a normal vector. Recall however, that we saw how to do this in the Cross Product section.

We can form the following two vectors from the given points.

$$\overrightarrow{PQ} = \langle 2,3,4\rangle \qquad\qquad \overrightarrow{PR} = \langle -1,1,2\rangle$$

These two vectors will lie completely in the plane since we formed them from points that were in the plane. Notice as well that there are many possible vectors to use here, we just chose two of the possibilities.

Now, we know that the cross product of two vectors will be orthogonal to both of these vectors. Since both of these are in the plane any vector that is orthogonal to both of these will also be orthogonal to the plane. Therefore, we can use the cross product as the normal vector.

$$\vec{n} = \overrightarrow{PQ}\times\overrightarrow{PR} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & 3 & 4 \\ -1 & 1 & 2 \end{vmatrix}\begin{matrix} \vec{i} & \vec{j} \\ 2 & 3 \\ -1 & 1 \end{matrix} = 2\vec{i} - 8\vec{j} + 5\vec{k}$$

The equation of the plane is then,

$$2\left(x-1\right)-8\left(y+2\right)+5\left(z-0\right)=0$$
$$2x-8y+5z=18$$

We used *P* for the point, but could have used any of the three points.

Example 2 Determine if the plane given by $-x+2z=10$ and the line given by $\vec{r}=\left\langle 5,2-t,10+4t\right\rangle$ are orthogonal, parallel or neither.

Solution

This is not as difficult a problem as it may at first appear to be. We can pick off a vector that is normal to the plane. This is $\vec{n}=\left\langle -1,0,2\right\rangle$. We can also get a vector that is parallel to the line. This is $v=\left\langle 0,-1,4\right\rangle$.

Now, if these two vectors are parallel then the line and the plane will be orthogonal. If you think about it this makes some sense. If $\vec{n}$ and $\vec{v}$ are parallel, then $\vec{v}$ is orthogonal to the plane, but $\vec{v}$ is also parallel to the line. So, if the two vectors are parallel the line and plane will be orthogonal.

Let's check this.

$$\vec{n}\times\vec{v}=\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -1 & 0 & 2 \\ 0 & -1 & 4 \end{vmatrix}\begin{matrix} \vec{i} & \vec{j} \\ -1 & 0 \\ 0 & -1 \end{matrix}=2\vec{i}+4\vec{j}+\vec{k}\neq\vec{0}$$

So, the vectors aren't parallel and so the plane and the line are not orthogonal.

Now, let's check to see if the plane and line are parallel. If the line is parallel to the plane then any vector parallel to the line will be orthogonal to the normal vector of the plane. In other words, if $\vec{n}$ and $\vec{v}$ are orthogonal then the line and the plane will be parallel.

Let's check this.

$$\vec{n}\bullet\vec{v}=0+0+8=8\neq 0$$

The two vectors aren't orthogonal and so the line and plane aren't parallel.

So, the line and the plane are neither orthogonal nor parallel.

Quadric Surfaces

In the previous two sections we've looked at lines and planes in three dimensions (or $\mathbb{R}^3$) and while these are used quite heavily at times in a Calculus class there are many other surfaces that are also used fairly regularly and so we need to take a look at those.

In this section we are going to be looking at quadric surfaces. Quadric surfaces are the graphs of any equation that can be put into the general form

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0$$

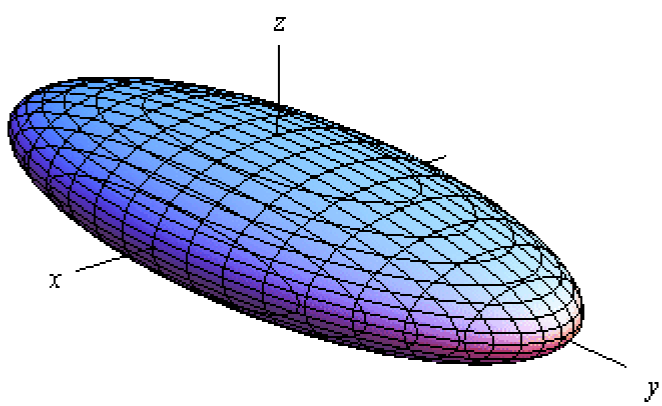
where *A*, … , *J* are constants.

There is no way that we can possibly list all of them, but there are some standard equations so here is a list of some of the more common quadric surfaces.

Ellipsoid
Here is the general equation of an ellipsoid.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Here is a sketch of a typical ellipsoid.



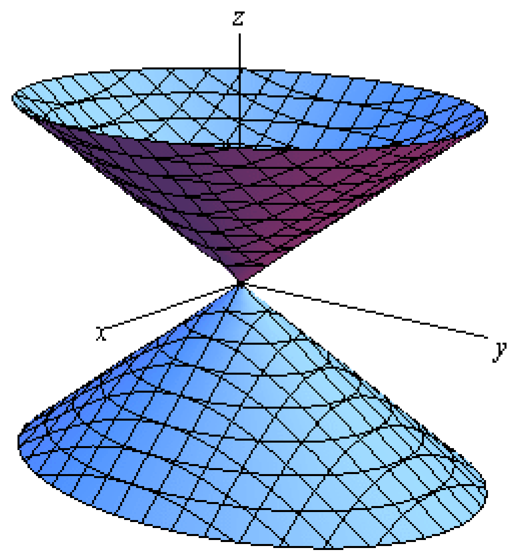
If $a = b = c$ then we will have a sphere.

Notice that we only gave the equation for the ellipsoid that has been centered on the origin. Clearly ellipsoids don't have to be centered on the origin. However, in order to make the discussion in this section a little easier we have chosen to concentrate on surfaces that are "centered" on the origin in one way or another.

Cone
Here is the general equation of a cone.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$$

Here is a sketch of a typical cone.



Note that this is the equation of a cone that will open along the *z*-axis. To get the equation of a cone that opens along one of the other axes all we need to do is make a slight modification of the equation. This will be the case for the rest of the surfaces that we'll be looking at in this section as well.

In the case of a cone the variable that sits by itself on one side of the equal sign will determine the axis that the cone opens up along. For instance, a cone that opens up along the *x*-axis will have the equation,

$$\frac{y^2}{b^2}+\frac{z^2}{c^2}=\frac{x^2}{a^2}$$

For most of the following surfaces we will not give the other possible formulas. We will however acknowledge how each formula needs to be changed to get a change of orientation for the surface.

Cylinder

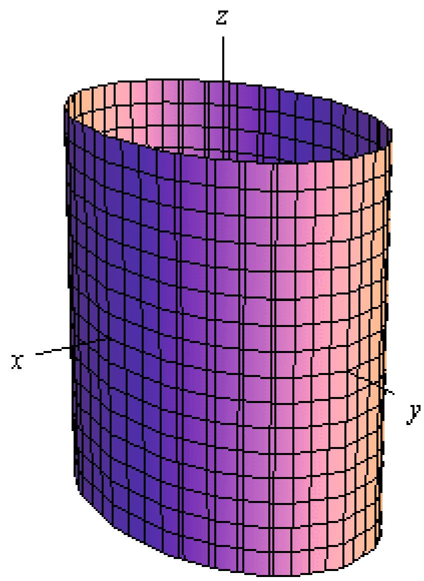
Here is the general equation of a cylinder.

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

This is a cylinder whose cross section is an ellipse. If $a=b$ we have a cylinder whose cross section is a circle. We'll be dealing with those kinds of cylinders more than the general form so the equation of a cylinder with a circular cross section is,

$$x^2+y^2=r^2$$

Here is a sketch of typical cylinder with an ellipse cross section.



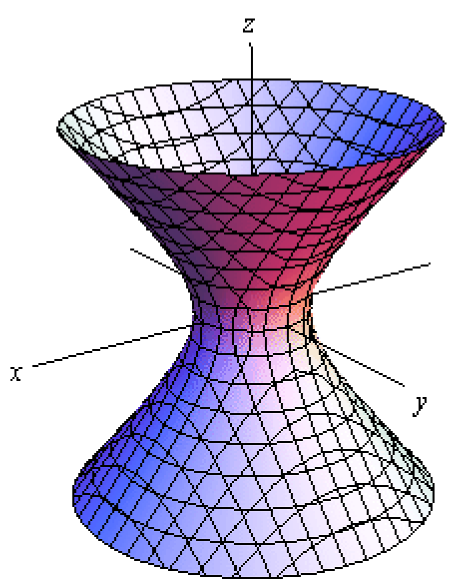
The cylinder will be centered on the axis corresponding to the variable that does not appear in the equation.

Be careful to not confuse this with a circle. In two dimensions it is a circle, but in three dimensions it is a cylinder.

Hyperboloid of One Sheet
Here is the equation of a hyperboloid of one sheet.

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

Here is a sketch of a typical hyperboloid of one sheet.

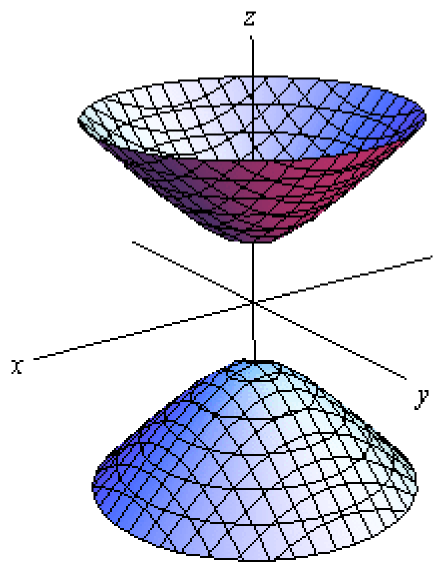


The variable with the negative in front of it will give the axis along which the graph is centered.

Hyperboloid of Two Sheets
Here is the equation of a hyperboloid of two sheets.

$$-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

Here is a sketch of a typical hyperboloid of two sheets.



The variable with the positive in front of it will give the axis along which the graph is centered.

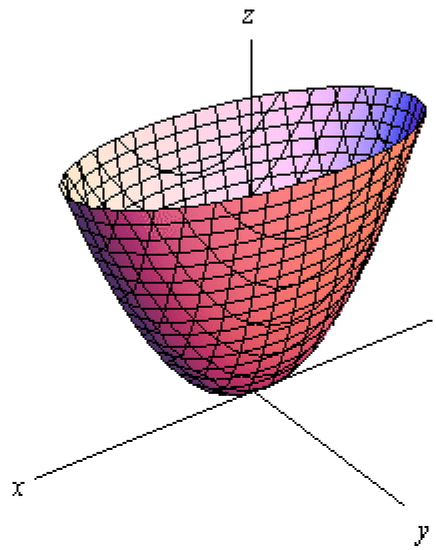
Notice that the only difference between the hyperboloid of one sheet and the hyperboloid of two sheets is the signs in front of the variables. They are exactly the opposite signs.

Elliptic Paraboloid
Here is the equation of an elliptic paraboloid.

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{z}{c}$$

As with cylinders this has a cross section of an ellipse and if $a=b$ it will have a cross section of a circle. When we deal with these we'll generally be dealing with the kind that have a circle for a cross section.

Here is a sketch of a typical elliptic paraboloid.

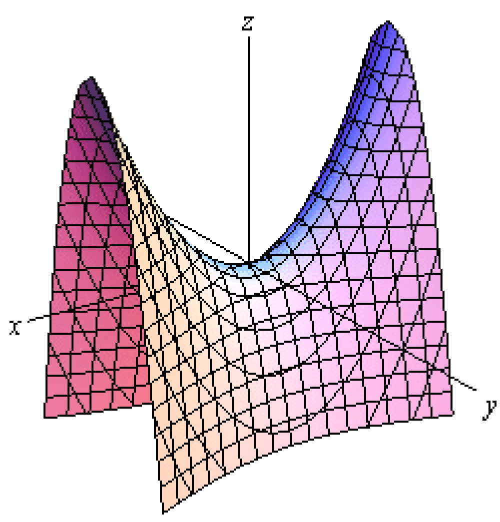


In this case the variable that isn't squared determines the axis upon which the paraboloid opens up. Also, the sign of c will determine the direction that the paraboloid opens. If c is positive then it opens up and if c is negative then it opens down.

Hyperbolic Paraboloid
Here is the equation of a hyperbolic paraboloid.

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c}$$

Here is a sketch of a typical hyperbolic paraboloid.



These graphs are vaguely saddle shaped and as with the elliptic paraboloid the sign of c will determine the direction in which the surface "opens up". The graph above is shown for c positive.

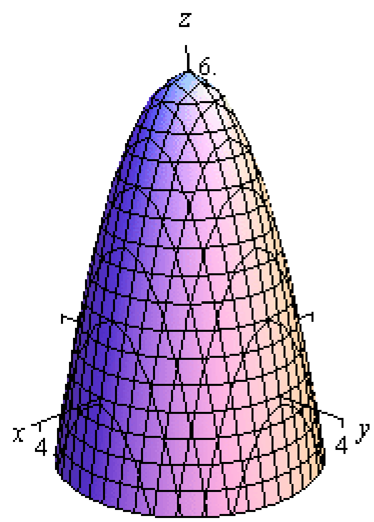
With the both of the types of paraboloids discussed above the surface can be easily moved up or down by adding/subtracting a constant from the left side.

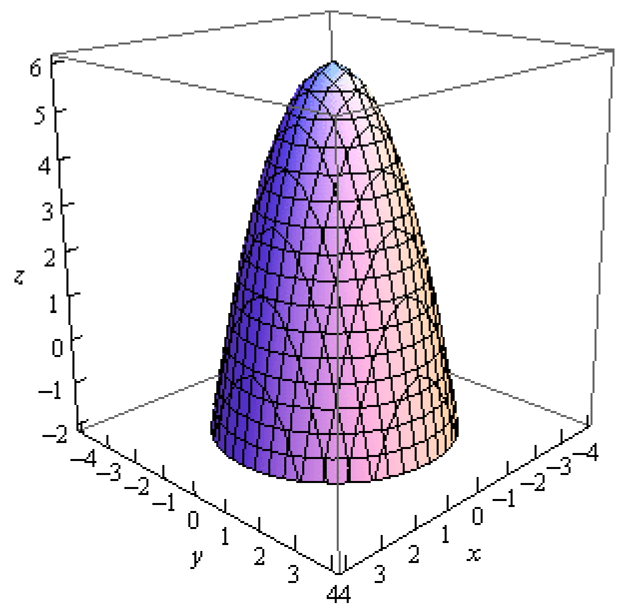
For instance

$$z = -x^2 - y^2 + 6$$

is an elliptic paraboloid that opens downward (be careful, the "-" is on the *x* and *y* instead of the *z*) and starts at $z = 6$ instead of $z = 0$.

Here are a couple of quick sketches of this surface.



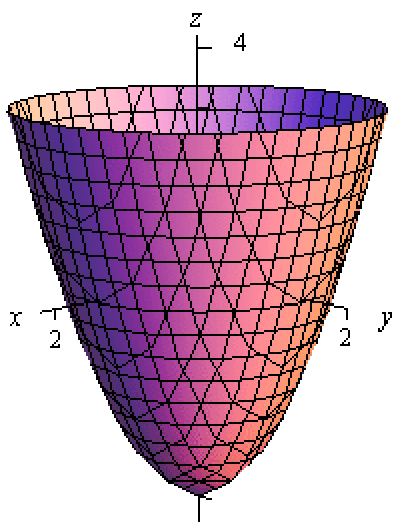


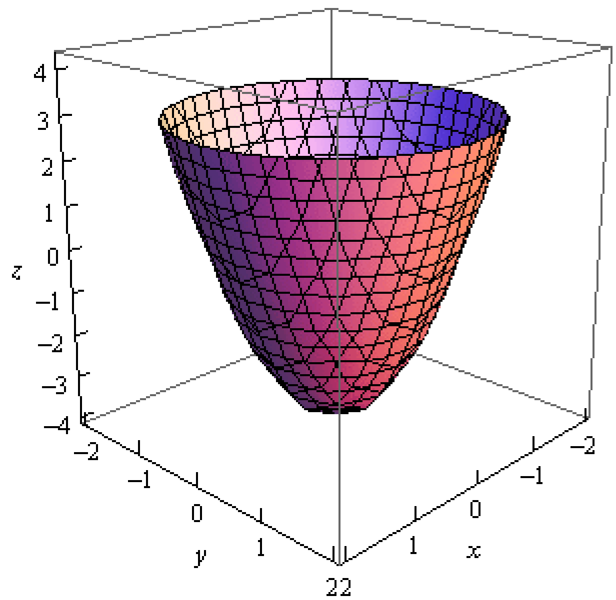
Note that we've given two forms of the sketch here. The sketch on the left has the standard set of axes but it is difficult to see the numbers on the axis. The sketch on the right has been "boxed" and this makes it easier to see the numbers to give a sense of perspective to the sketch. In most sketches that actually involve numbers on the axis system we will give both sketches to help get a feel for what the sketch looks like.

Functions of Several Variables

In this section we want to go over some of the basic ideas about functions of more than one variable.

First, remember that graphs of functions of two variables, $z = f(x, y)$ are surfaces in three dimensional space. For example here is the graph of $z = 2x^2 + 2y^2 - 4$.





This is an elliptic parabaloid and is an example of a quadric surface. We saw several of these in the previous section. We will be seeing quadric surfaces fairly regularly later on in Calculus III.

Another common graph that we'll be seeing quite a bit in this course is the graph of a plane. We have a convention for graphing planes that will make them a little easier to graph and hopefully visualize.

Recall that the equation of a plane is given by

$$ax + by + cz = d$$

or if we solve this for z we can write it in terms of function notation. This gives,

$$f(x, y) = Ax + By + D$$

To graph a plane we will generally find the intersection points with the three axes and then graph the triangle that connects those three points. This triangle will be a portion of the plane and it will give us a fairly decent idea on what the plane itself should look like. For example let's graph the plane given by,

$$f(x, y) = 12 - 3x - 4y$$

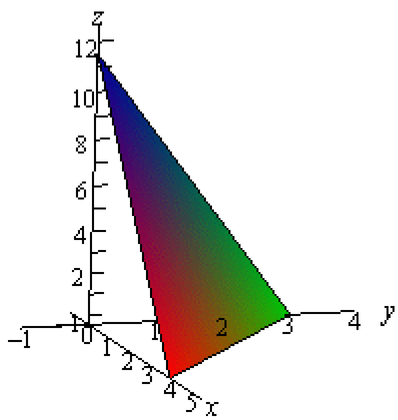
For purposes of graphing this it would probably be easier to write this as,

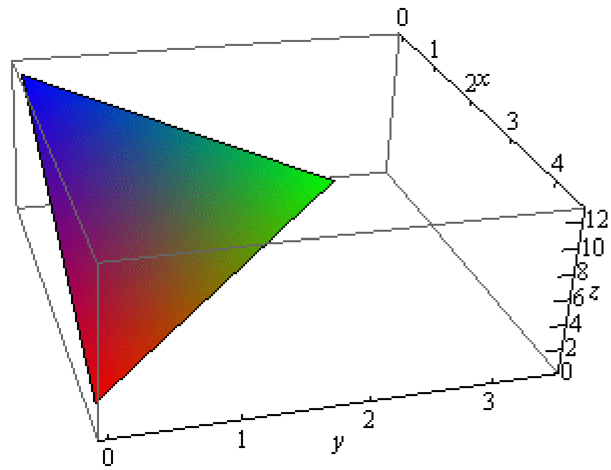
$$z = 12 - 3x - 4y \qquad \Rightarrow \qquad 3x + 4y + z = 12$$

Now, each of the intersection points with the three main coordinate axes is defined by the fact that two of the coordinates are zero. For instance, the intersection with the *z*-axis is defined by $x = y = 0$. So, the three intersection points are,

$$x-\text{axis}: \ (4,0,0)$$
$$y-\text{axis}: \ (0,3,0)$$
$$z-\text{axis}: \ (0,0,12)$$

Here is the graph of the plane.





Now, to extend this out, graphs of functions of the form $w = f(x, y, z)$ would be four dimensional surfaces. Of course we can't graph them, but it doesn't hurt to point this out.

We next want to talk about the domains of functions of more than one variable. Recall that domains of functions of a single variable, $y = f(x)$, consisted of all the values of *x* that we could plug into the function and get back a real number. Now, if we think about it, this means that the domain of a function of a single variable is an interval (or intervals) of values from the number line, or one dimensional space.

The domain of functions of two variables, $z = f(x, y)$, are regions from two dimensional space and consist of all the coordinate pairs, (x, y), that we could plug into the function and get back a real number.

Example 1 Determine the domain of each of the following.

(a) $f(x,y)=\sqrt{x+y}$ [Solution]

(b) $f(x,y)=\sqrt{x}+\sqrt{y}$ [Solution]

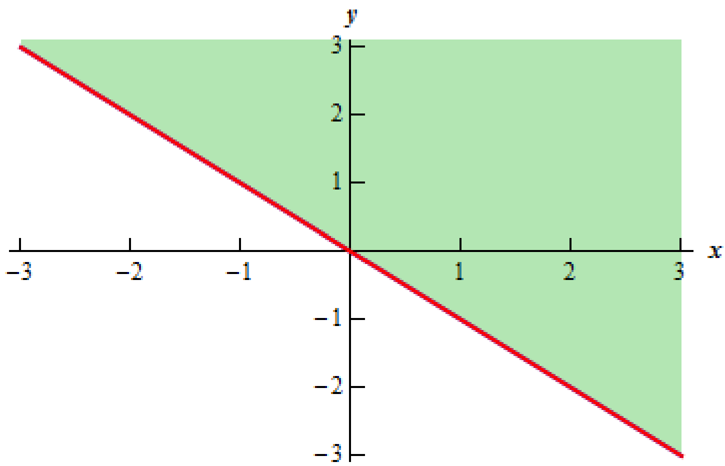
(c) $f(x,y)=\ln\left(9-x^2-9y^2\right)$ [Solution]

Solution

(a) In this case we know that we can't take the square root of a negative number so this means that we must require,

$$x+y\geq 0$$

Here is a sketch of the graph of this region.

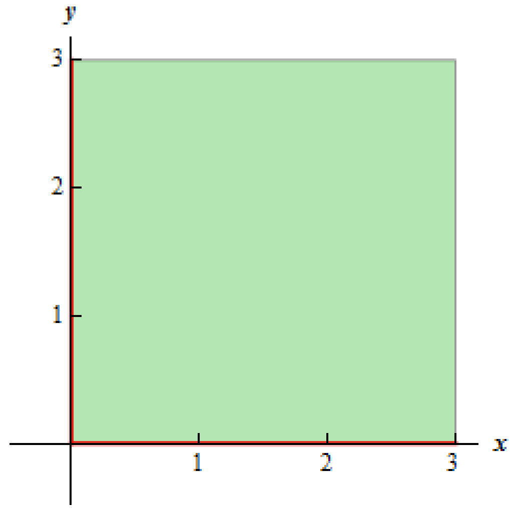


[Return to Problems]

(b) This function is different from the function in the previous part. Here we must require that,

$$x\geq 0 \quad \text{and} \quad y\geq 0$$

and they really do need to be separate inequalities. There is one for each square root in the function. Here is the sketch of this region.

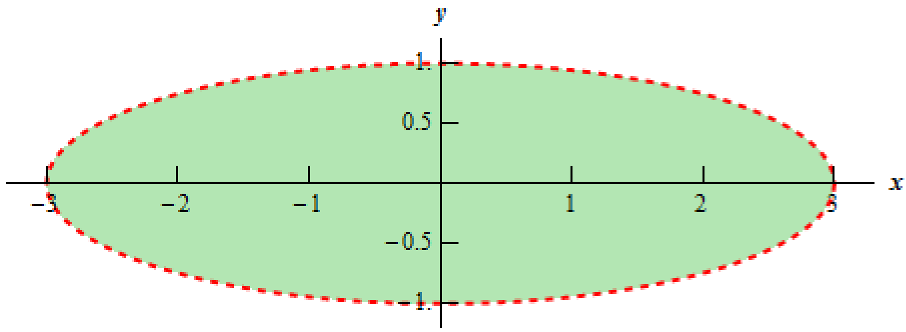


[Return to Problems]

(c) In this final part we know that we can't take the logarithm of a negative number or zero. Therefore we need to require that,

$$9 - x^2 - 9y^2 > 0 \qquad \Rightarrow \qquad \frac{x^2}{9} + y^2 < 1$$

and upon rearranging we see that we need to stay interior to an ellipse for this function. Here is a sketch of this region.



[Return to Problems]

Note that domains of functions of three variables, $w = f(x, y, z)$, will be regions in three dimensional space.

Example 2 Determine the domain of the following function,

$$f(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2 - 16}}$$

Solution

In this case we have to deal with the square root and division by zero issues. These will require,

$$x^2 + y^2 + z^2 - 16 > 0 \qquad \Rightarrow \qquad x^2 + y^2 + z^2 > 16$$

So, the domain for this function is the set of points that lies completely outside a sphere of radius 4 centered at the origin.

The next topic that we should look at is that of **level curves** or **contour curves**. The level curves of the function $z = f(x, y)$ are two dimensional curves we get by setting $z = k$, where k is any number. So the equations of the level curves are $f(x, y) = k$. Note that sometimes the equation will be in the form $f(x, y, z) = 0$ and in these cases the equations of the level curves are $f(x, y, k) = 0$.

You've probably seen level curves (or contour curves, whatever you want to call them) before. If you've ever seen the elevation map for a piece of land, this is nothing more than the contour curves for the function that gives the elevation of the land in that area. Of course, we probably don't have the function that gives the elevation, but we can at least graph the contour curves.

Let's do a quick example of this.

Example 3 Identify the level curves of $f(x, y) = \sqrt{x^2 + y^2}$. Sketch a few of them.

Solution

First, for the sake of practice, let's identify what this surface given by $f(x, y)$ is. To do this let's rewrite it as,

$$z = \sqrt{x^2 + y^2}$$

Now, this equation is not listed in the Quadric Surfaces section, but if we square both sides we get,

$$z^2 = x^2 + y^2$$

and this is listed in that section. So, we have a cone, or at least a portion of a cone. Since we know that square roots will only return positive numbers, it looks like we've only got the upper half of a cone.

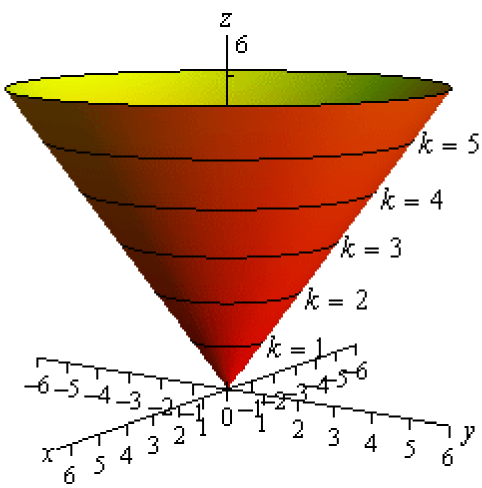
Note that this was not required for this problem. It was done for the practice of identifying the surface and this may come in handy down the road.

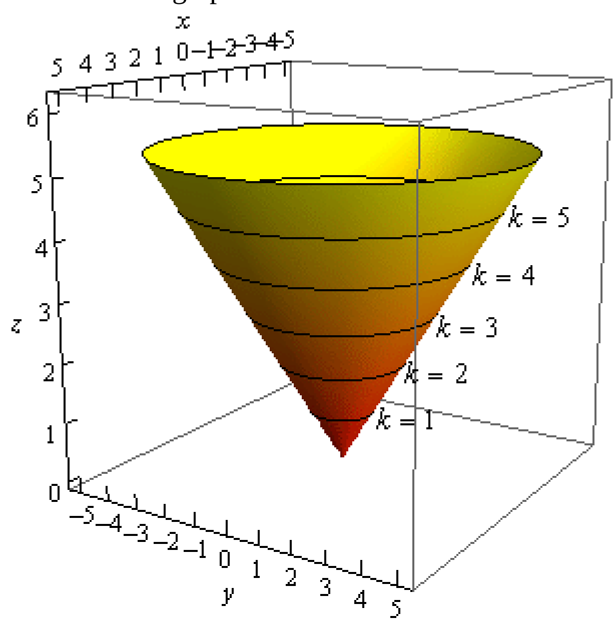
Now on to the real problem. The level curves (or contour curves) for this surface are given by the equation are found by substituting $z = k$. In the case of our example this is,

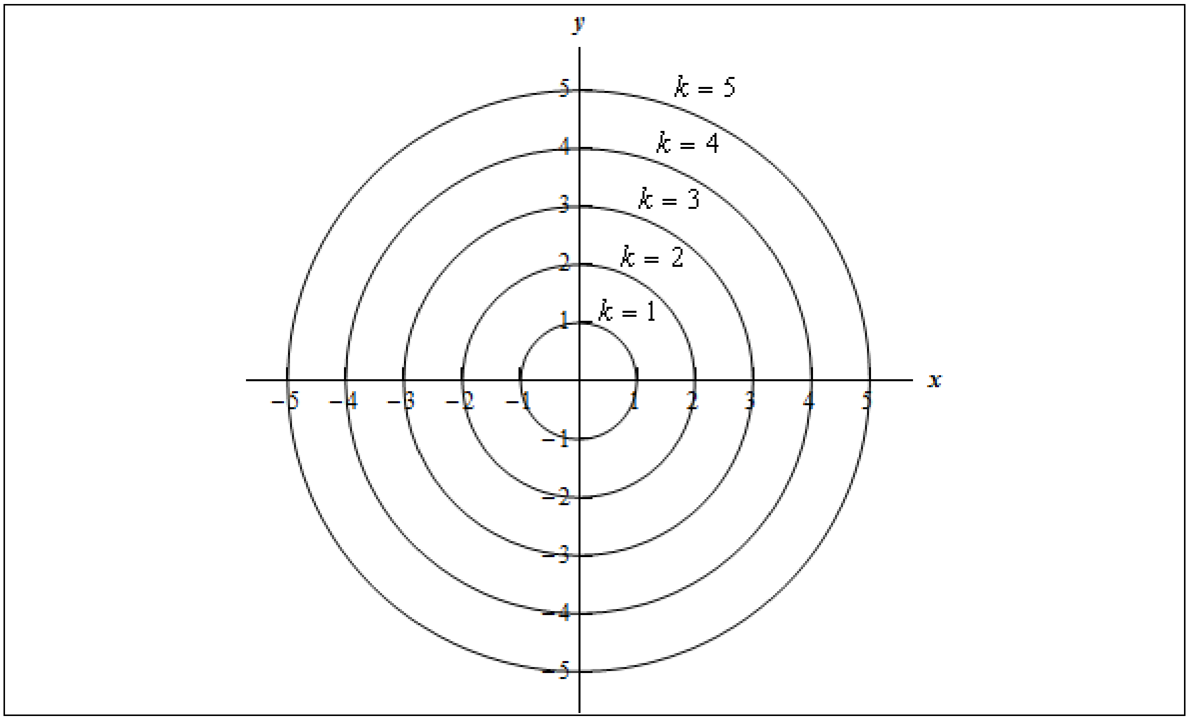
$$k = \sqrt{x^2 + y^2} \qquad \Rightarrow \qquad x^2 + y^2 = k^2$$

where k is any number. So, in this case, the level curves are circles of radius k with center at the origin.

We can graph these in one of two ways. We can either graph them on the surface itself or we can graph them in a two dimensional axis system. Here is each graph for some values of k.







Note that we can think of contours in terms of the intersection of the surface that is given by $z = f(x, y)$ and the plane $z = k$. The contour will represent the intersection of the surface and the plane.

For functions of the form $f(x, y, z)$ we will occasionally look at **level surfaces**. The equations of level surfaces are given by $f(x, y, z) = k$ where *k* is any number.

The final topic in this section is that of **traces**. In some ways these are similar to contours. As noted above we can think of contours as the intersection of the surface given by $z = f(x, y)$ and the plane $z = k$. Traces of surfaces are curves that represent the intersection of the surface and the plane given by $x = a$ or $y = b$.

Let's take a quick look at an example of traces.

Example 4 Sketch the traces of $f(x, y) = 10 - 4x^2 - y^2$ for the plane $x = 1$ and $y = 2$.

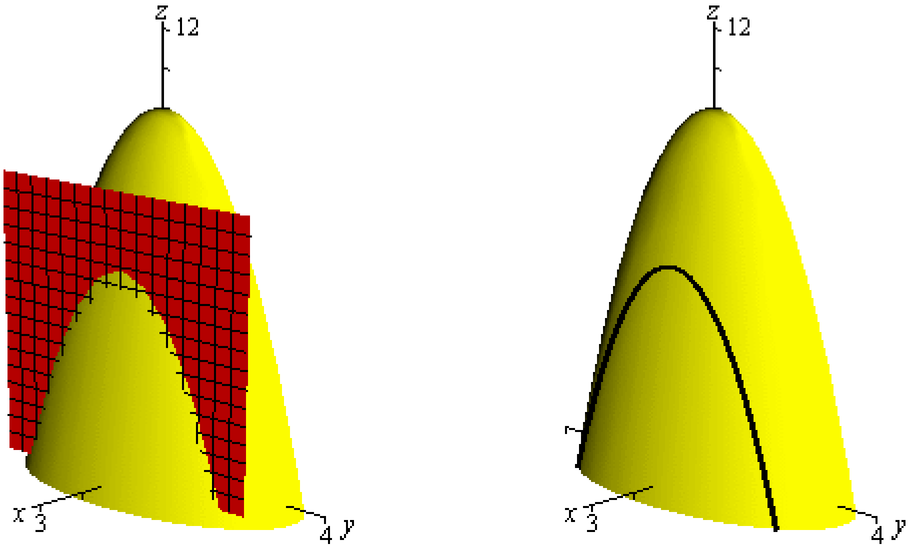
Solution

We'll start with $x = 1$. We can get an equation for the trace by plugging $x = 1$ into the equation. Doing this gives,

$$z = f(1, y) = 10 - 4(1)^2 - y^2 \qquad \Rightarrow \qquad z = 6 - y^2$$

and this will be graphed in the plane given by $x = 1$.

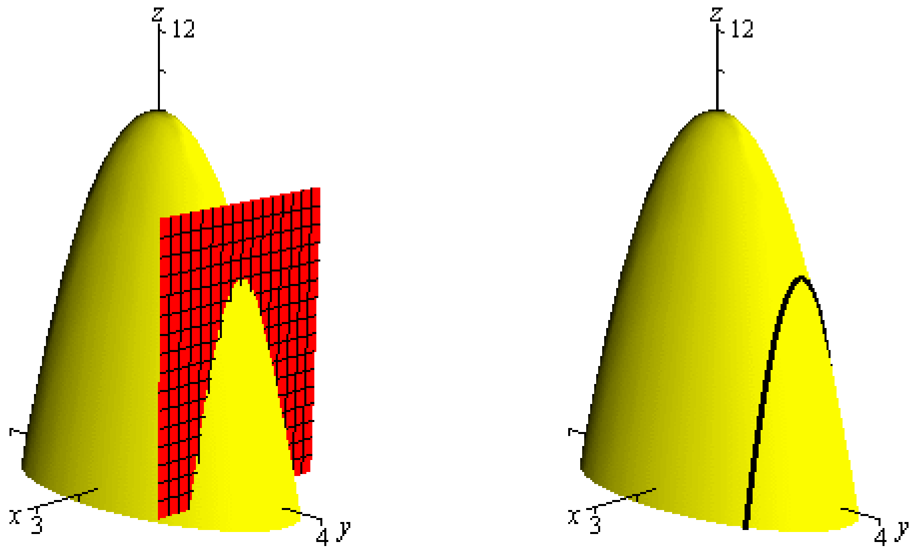
Below are two graphs. The graph on the left is a graph showing the intersection of the surface and the plane given by $x = 1$. On the right is a graph of the surface and the trace that we are after in this part.



For $y = 2$ we will do pretty much the same thing that we did with the first part. Here is the equation of the trace,

$$z = f(x,2) = 10 - 4x^2 - (2)^2 \qquad \Rightarrow \qquad z = 6 - 4x^2$$

and here are the sketches for this case.



Vector Functions

We first saw vector functions back when we were looking at the Equation of Lines. In that section we talked about them because we wrote down the equation of a line in $\mathbb{R}^3$ in terms of a **vector function** (sometimes called a **vector-valued function**). In this section we want to look a little closer at them and we also want to look at some vector functions in $\mathbb{R}^3$ other than lines.

A vector function is a function that takes one or more variables and returns a vector. We'll spend most of this section looking at vector functions of a single variable as most of the places where vector functions show up here will be vector functions of single variables. We will however briefly look at vector functions of two variables at the end of this section.

A vector functions of a single variable in $\mathbb{R}^2$ and $\mathbb{R}^3$ have the form,

$$\vec{r}(t) = \langle f(t), g(t) \rangle \qquad\qquad \vec{r}(t) = \langle f(t), g(t), h(t) \rangle$$

respectively, where $f(t)$, $g(t)$ and $h(t)$ are called the **component functions**.

The main idea that we want to discuss in this section is that of graphing and identifying the graph given by a vector function. Before we do that however, we should talk briefly about the domain of a vector function. The **domain** of a vector function is the set of all *t*'s for which all the component functions are defined.

Example 1 Determine the domain of the following function.

$$\vec{r}(t) = \left\langle \cos t, \ln(4-t), \sqrt{t+1} \right\rangle$$

Solution

The first component is defined for all *t*'s. The second component is only defined for $t < 4$. The third component is only defined for $t \geq -1$. Putting all of these together gives the following domain.

$$[-1, 4)$$

This is the largest possible interval for which all three components are defined.

Let's now move into looking at the graph of vector functions. In order to graph a vector function all we do is think of the vector returned by the vector function as a position vector for points on the graph. Recall that a position vector, say $\vec{v} = \langle a, b, c \rangle$, is a vector that starts at the origin and ends at the point (a, b, c).

So, in order to sketch the graph of a vector function all we need to do is plug in some values of *t* and then plot points that correspond to the resulting position vector we get out of the vector function.

Because it is a little easier to visualize things we'll start off by looking at graphs of vector functions in $\mathbb{R}^2$.

Example 2 Sketch the graph of each of the following vector functions.

(a) $\vec{r}(t) = \langle t, 1 \rangle$ [Solution]

(b) $\vec{r}(t) = \langle t, t^3 - 10t + 7 \rangle$ [Solution]

Solution

(a) $\vec{r}(t) = \langle t, 1 \rangle$

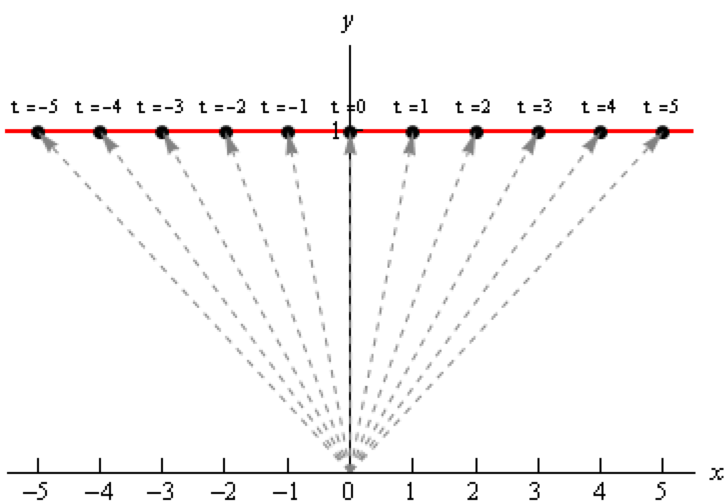
Okay, the first thing that we need to do is plug in a few values of *t* and get some position vectors. Here are a few,

$$\vec{r}(-3) = \langle -3, 1 \rangle \qquad \vec{r}(-1) = \langle -1, 1 \rangle \qquad \vec{r}(2) = \langle 2, 1 \rangle \qquad \vec{r}(5) = \langle 5, 1 \rangle$$

So, what this tells us is that the following points are all on the graph of this vector function.

$$(-3,1) \qquad (-1,1) \qquad (2,1) \qquad (5,1)$$

Here is a sketch of this vector function.



In this sketch we've included many more evaluations that just those above. Also note that we've put in the position vectors (in gray and dashed) so you can see how all this is working. Note however, that in practice the position vectors are generally not included in the sketch.

In this case it looks like we've got the graph of the line $y = 1$.

[Return to Problems]

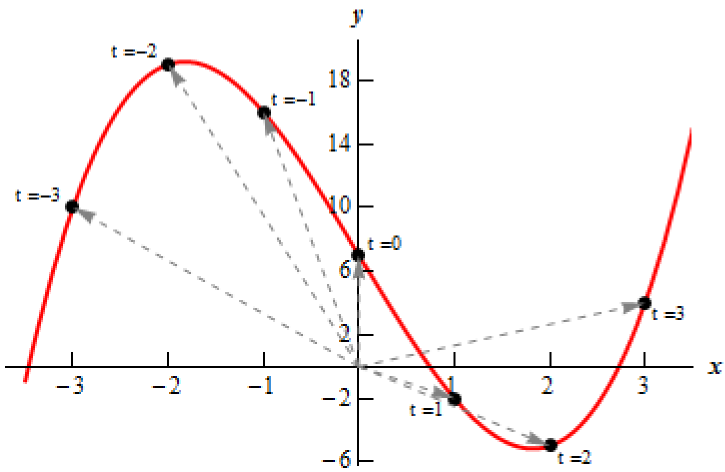
(b) $\vec{r}(t) = \langle t, t^3 - 10t + 7 \rangle$

Here are a couple of evaluations for this vector function.

$$\vec{r}(-3) = \langle -3, 10 \rangle \qquad \vec{r}(-1) = \langle -1, 16 \rangle \qquad \vec{r}(1) = \langle 1, -2 \rangle \qquad \vec{r}(3) = \langle 3, 4 \rangle$$

So, we've got a few points on the graph of this function. However, unlike the first part this isn't really going to be enough points to get a good idea of this graph. In general, it can take quite a few function evaluations to get an idea of what the graph is and it's usually easier to use a computer to do the graphing.

Here is a sketch of this graph. We've put in a few vectors/evaluations to illustrate them, but the reality is that we did have to use a computer to get a good sketch here.



[Return to Problems]

Both of the vector functions in the above example were in the form,

$$\vec{r}(t) = \langle t, g(t) \rangle$$

and what we were really sketching is the graph of $y = g(x)$ as you probably caught onto. Let's graph a couple of other vector functions that do not fall into this pattern.

Example 3 Sketch the graph of each of the following vector functions.

(a) $\vec{r}(t) = \langle 6\cos t, 3\sin t \rangle$ [Solution]

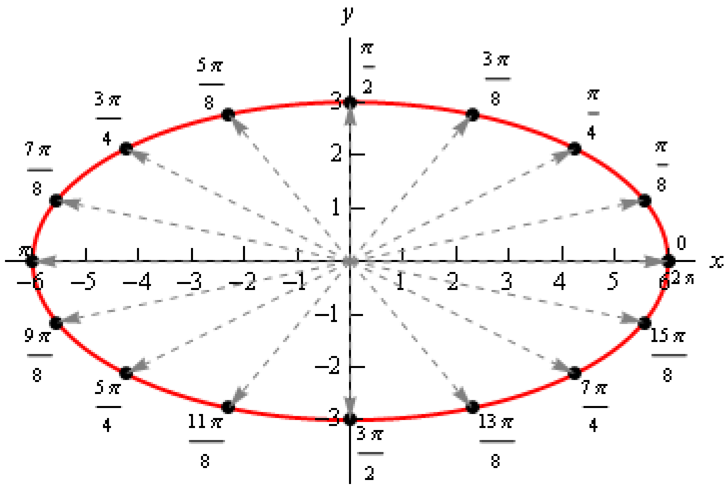
(b) $\vec{r}(t) = \langle t - 2\sin t, t^2 \rangle$ [Solution]

Solution

As we saw in the last part of the previous example it can really take quite a few function evaluations to really be able to sketch the graph of a vector function. Because of that we'll be skipping all the function evaluations here and just giving the graph. The main point behind this set of examples is to not get you too locked into the form we were looking at above. The first part will also lead to an important idea that we'll discuss after this example.

So, with that said here are the sketches of each of these.

(a) $\vec{r}(t) = \langle 6\cos t, 3\sin t \rangle$

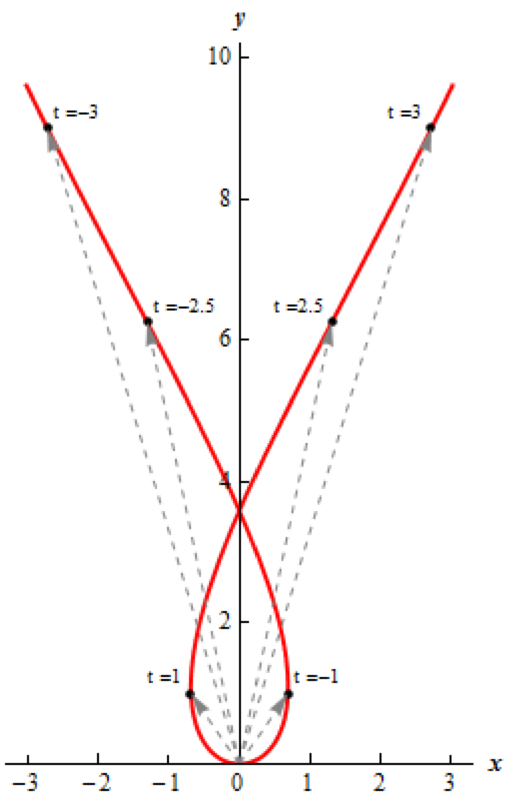


So, in this case it looks like we've got an ellipse.

[Return to Problems]

(b) $\vec{r}(t) = \langle t - 2\sin t, t^2 \rangle$

Here's the sketch for this vector function.



[Return to Problems]

Before we move on to vector functions in $\mathbb{R}^3$ let's go back and take a quick look at the first vector function we sketched in the previous example, $\vec{r}(t) = \langle 6\cos t, 3\sin t \rangle$. The fact that we got an ellipse here should not come as a surprise to you. We know that the first component function gives the *x* coordinate and the second component function gives the *y* coordinates of the point that we graph. If we strip these out to make this clear we get,

$$x = 6\cos t \qquad\qquad y = 3\sin t$$

This should look familiar to you. Back when we were looking at Parametric Equations we saw that this was nothing more than one of the sets of parametric equations that gave an ellipse.

This is an important idea in the study of vector functions. Any vector function can be broken down into a set of parametric equations that represent the same graph. In general, the two dimensional vector function, $\vec{r}\left(t\right) = \left\langle f\left(t\right), g\left(t\right)\right\rangle$, can be broken down into the parametric equations,

$$x = f\left(t\right) \qquad\qquad y = g\left(t\right)$$

Likewise, a three dimensional vector function, $\vec{r}\left(t\right) = \left\langle f\left(t\right), g\left(t\right), h\left(t\right)\right\rangle$, can be broken down into the parametric equations,

$$x = f\left(t\right) \qquad y = g\left(t\right) \qquad z = h\left(t\right)$$

Do not get too excited about the fact that we're now looking at parametric equations in $\mathbb{R}^3$. They work in exactly the same manner as parametric equations in $\mathbb{R}^2$ which we're used to dealing with already. The only difference is that we now have a third component.

Let's take a look at a couple of graphs of vector functions.

Example 4 Sketch the graph of the following vector function.

$$\vec{r}\left(t\right) = \left\langle 2-4t, -1+5t, 3+t\right\rangle$$

Solution

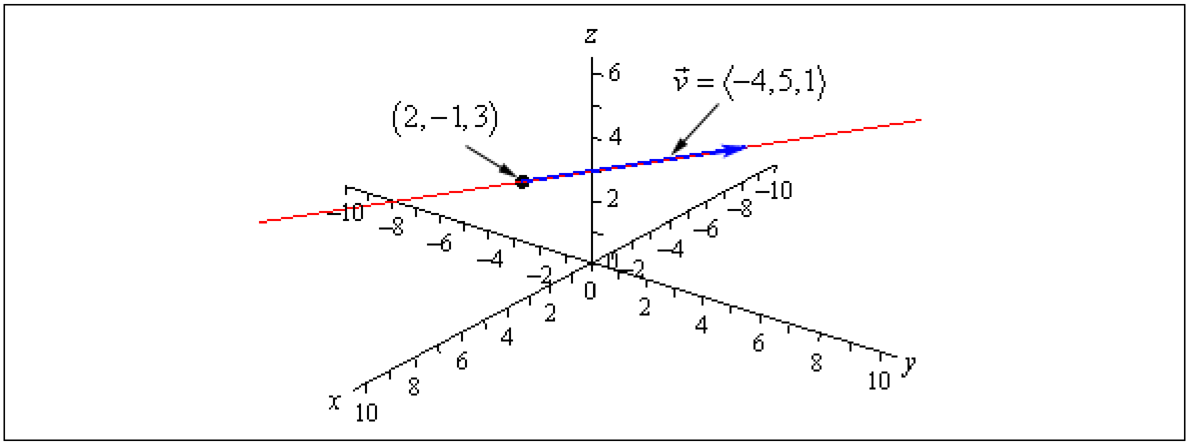
Notice that this is nothing more than a line. It might help if we rewrite it a little.

$$\vec{r}\left(t\right) = \left\langle 2,-1,3\right\rangle + t\left\langle -4,5,1\right\rangle$$

In this form we can see that this is the equation of a line that goes through the point $\left(2,-1,3\right)$ and is parallel to the vector $\vec{v} = \left\langle -4,5,1\right\rangle$.

To graph this line all that we need to do is plot the point and then sketch in the parallel vector. In order to get the sketch will assume that the vector is on the line and will start at the point in the line. To sketch in the line all we do this is extend the parallel vector into a line.

Here is a sketch.



Example 5 Sketch the graph of the following vector function.

$$\vec{r}(t) = \langle 2\cos t, 2\sin t, 3\rangle$$

Solution

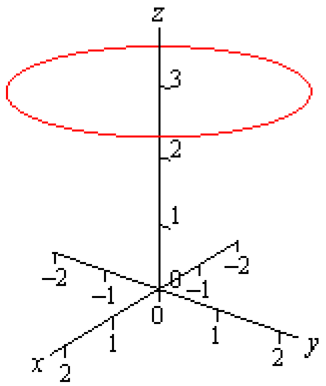
In this case to see what we've got for a graph let's get the parametric equations for the curve.

$$x = 2\cos t \qquad\qquad y = 2\sin t \qquad\qquad z = 3$$

If we ignore the *z* equation for a bit we'll recall (hopefully) that the parametric equations for *x* and *y* give a circle of radius 2 centered on the origin (or about the *z*-axis since we are in $\mathbb{R}^3$).

Now, all the parametric equations here tell us is that no matter what is going on in the graph all the *z* coordinates must be 3. So, we get a circle of radius 2 centered on the *z*-axis and at the level of $z = 3$.

Here is a sketch.



Note that it is very easy to modify the above vector function to get a circle centered on the *x* or *y*-axis as well. For instance,

$$\vec{r}(t) = \langle 10\sin t, -3, 10\cos t\rangle$$

will be a circle of radius 10 centered on the *y*-axis and at $y = -3$. In other words, as long as two of the terms are a sine and a cosine (with the same coefficient) and the other is a fixed number then we will have a circle that is centered on the axis that is given by the fixed number.

Let's take a look at a modification of this.

Example 6 Sketch the graph of the following vector function.

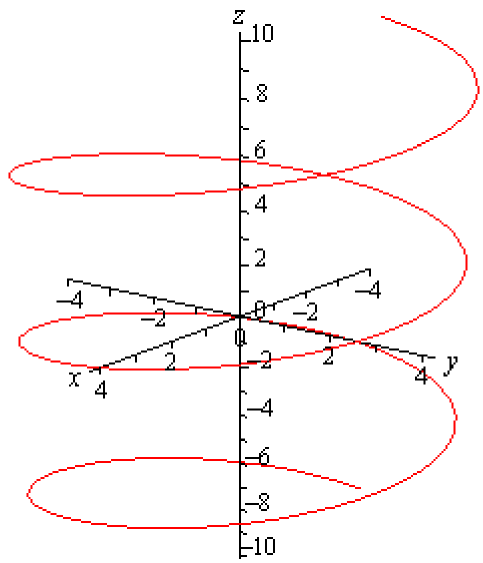
$$\vec{r}(t) = \langle 4\cos t, 4\sin t, t \rangle$$

Solution

If this one had a constant in the *z* component we would have another circle. However, in this case we don't have a constant. Instead we've got a *t* and that will change the curve. However, because the *x* and *y* component functions are still a circle in parametric equations our curve should have a circular nature to it in some way.

In fact, the only change is in the *z* component and as *t* increases the *z* coordinate will increase. Also, as *t* increases the *x* and *y* coordinates will continue to form a circle centered on the *z*-axis. Putting these two ideas together tells us that as we increase *t* the circle that is being traced out in the *x* and *y* directions should also be rising.

Here is a sketch of this curve.



So, we've got a helix (or spiral, depending on what you want to call it) here.

As with circles the component that has the *t* will determine the axis that the helix rotates about. For instance,

$$\vec{r}(t) = \langle t, 6\cos t, 6\sin t \rangle$$

is a helix that rotates around the *x*-axis.

Also note that if we allow the coefficients on the sine and cosine for both the circle and helix to be different we will get ellipses.

For example,

$$\vec{r}(t) = \langle 9\cos t, t, 2\sin t \rangle$$

will be a helix that rotates about the *y*-axis and is in the shape of an ellipse.

There is a nice formula that we should derive before moving onto vector functions of two variables.

Example 7 Determine the vector equation for the line segment starting at the point $P = (x_1, y_1, z_1)$ and ending at the point $Q = (x_2, y_2, z_2)$.

Solution
It is important to note here that we only want the equation of the line segment that starts at *P* and ends at *Q*. We don't want any other portion of the line and we do want the direction of the line segment preserved as we increase *t*. With all that said, let's not worry about that and just find the vector equation of the line that passes through the two points. Once we have this we will be able to get what we're after.

So, we need a point on the line. We've got two and we will use *P*. We need a vector that is parallel to the line and since we've got two points we can find the vector between them. This vector will lie on the line and hence be parallel to the line. Also, let's remember that we want to preserve the starting and ending point of the line segment so let's construct the vector using the same "orientation".

$$\vec{v} = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1 \rangle$$

Using this vector and the point *P* we get the following vector equation of the line.

$$\vec{r}(t) = \langle x_1, y_1, z_1 \rangle + t\langle x_2 - x_1, y_2 - y_1, z_2 - z_1 \rangle$$

While this is the vector equation of the line, let's rewrite the equation slightly.

$$\begin{aligned}\vec{r}(t) &= \langle x_1, y_1, z_1 \rangle + t\langle x_2, y_2, z_2 \rangle - t\langle x_1, y_1, z_1 \rangle \\ &= (1-t)\langle x_1, y_1, z_1 \rangle + t\langle x_2, y_2, z_2 \rangle\end{aligned}$$

This is the equation of the line that contains the points *P* and *Q*. We of course just want the line segment that starts at *P* and ends at *Q*. We can get this by simply restricting the values of *t*.

Notice that

$$\vec{r}(0) = \langle x_1, y_1, z_1 \rangle \qquad\qquad \vec{r}(1) = \langle x_2, y_2, z_2 \rangle$$

So, if we restrict *t* to be between zero and one we will cover the line segment and we will start and end at the correct point.

So the vector equation of the line segment that starts at $P = (x_1, y_1, z_1)$ and ends at $Q = (x_2, y_2, z_2)$ is,

$$\vec{r}(t) = (1-t)\langle x_1, y_1, z_1 \rangle + t\langle x_2, y_2, z_2 \rangle \qquad 0 \le t \le 1$$

As noted briefly at the beginning of this section we can also have vector functions of two variables. In these case the graphs of vector function of two variables are surfaces. So, to make sure that we don't forget that let's work an example with that as well.

Example 8 Identify the surface that is described by $\vec{r}\left(x,y\right)=x\,\vec{i}+y\,\vec{j}+\left(x^2+y^2\right)\vec{k}$.

Solution
First, notice that in this case the vector function will in fact be a function of two variables. This will always be the case when we are using vector functions to represent surfaces.

To identify the surface let's go back to parametric equations.

$$x=x \qquad\qquad y=y \qquad\qquad z=x^2+y^2$$

The first two are really only acknowledging that we are picking *x* and *y* for free and then determining *z* from our choices of these two. The last equation is the one that we want. We should recognize that function from the section on quadric surfaces. The third equation is the equation of an elliptic paraboloid and so the vector function represents an elliptic paraboloid.

As a final topic for this section let's generalize the idea from the previous example and note that given any function of one variable ($y=f\left(x\right)$ or $x=h\left(y\right)$) or any function of two variables ($z=g\left(x,y\right)$, $x=g\left(y,z\right)$, or $y=g\left(x,z\right)$) we can always write down a vector form of the equation.

For a function of one variable this will be,

$$\vec{r}\left(x\right)=x\,\vec{i}+f\left(x\right)\vec{j} \qquad\qquad \vec{r}\left(y\right)=h\left(y\right)\vec{i}+y\,\vec{j}$$

and for a function of two variables the vector form will be,

$$\vec{r}\left(x,y\right)=x\,\vec{i}+y\,\vec{j}+g\left(x,y\right)\vec{k} \qquad\qquad \vec{r}\left(y,z\right)=g\left(y,z\right)\vec{i}+y\,\vec{j}+z\,\vec{k}$$

$$\vec{r}\left(x,z\right)=x\,\vec{i}+g\left(x,z\right)\vec{j}+z\,\vec{k}$$

depending upon the original form of the function.

For example the hyperbolic paraboloid $y=2x^2-5z^2$ can be written as the following vector function.

$$\vec{r}\left(x,z\right)=x\,\vec{i}+\left(2x^2-5z^2\right)\vec{j}+z\,\vec{k}$$

This is a fairly important idea and we will be doing quite a bit of this kind of thing in Calculus III.

Calculus with Vector Functions

In this section we need to talk briefly about limits, derivatives and integrals of vector functions. As you will see, these behave in a fairly predictable manner. We will be doing all of the work in $\mathbb{R}^3$ but we can naturally extend the formulas/work in this section to $\mathbb{R}^n$ (*i.e.* *n*-dimensional space).

Let's start with limits. Here is the limit of a vector function.

$$\begin{aligned}\lim_{t\to a}\vec{r}(t) &= \lim_{t\to a}\left\langle f(t), g(t), h(t)\right\rangle \\ &= \left\langle \lim_{t\to a} f(t), \lim_{t\to a} g(t), \lim_{t\to a} h(t)\right\rangle \\ &= \lim_{t\to a} f(t)\,\vec{i} + \lim_{t\to a} g(t)\,\vec{j} + \lim_{t\to a} h(t)\,\vec{k}\end{aligned}$$

So, all that we do is take the limit of each of the component's functions and leave it as a vector.

Example 1 Compute $\lim\limits_{t\to 1}\vec{r}(t)$ where $\vec{r}(t) = \left\langle t^3, \dfrac{\sin(3t-3)}{t-1}, \mathbf{e}^{2t}\right\rangle$.

Solution

There really isn't all that much to do here.

$$\begin{aligned}\lim_{t\to 1}\vec{r}(t) &= \left\langle \lim_{t\to 1} t^3, \lim_{t\to 1}\frac{\sin(3t-3)}{t-1}, \lim_{t\to 1}\mathbf{e}^{2t}\right\rangle \\ &= \left\langle \lim_{t\to 1} t^3, \lim_{t\to 1}\frac{3\cos(3t-3)}{1}, \lim_{t\to 1}\mathbf{e}^{2t}\right\rangle \\ &= \left\langle 1, 3, \mathbf{e}^2\right\rangle\end{aligned}$$

Notice that we had to use L'Hospital's Rule on the *y* component.

Now let's take care of derivatives and after seeing how limits work it shouldn't be too surprising that we have the following for derivatives.

$$\vec{r}\,'(t) = \left\langle f'(t), g'(t), h'(t)\right\rangle = f'(t)\,\vec{i} + g'(t)\,\vec{j} + h'(t)\,\vec{k}$$

Example 2 Compute $\vec{r}\,'(t)$ for $\vec{r}(t) = t^6\,\vec{i} + \sin(2t)\,\vec{j} - \ln(t+1)\,\vec{k}$.

Solution

There really isn't too much to this problem other than taking the derivatives.

$$\vec{r}\,'(t) = 6t^5\,\vec{i} + 2\cos(2t)\,\vec{j} - \frac{1}{t+1}\vec{k}$$

Most of the basic facts that we know about derivatives still hold however, just to make it clear here are some facts about derivatives of vector functions.

Facts

$$\frac{d}{dt}\left(\vec{u}+\vec{v}\right)=\vec{u}'+\vec{v}'$$

$$\left(c\vec{u}\right)'=c\,\vec{u}'$$

$$\frac{d}{dt}\left(f\left(t\right)\vec{u}\left(t\right)\right)=f'\left(t\right)\vec{u}\left(t\right)+f\left(t\right)\vec{u}'$$

$$\frac{d}{dt}\left(\vec{u}\bullet\vec{v}\right)=\vec{u}'\bullet\vec{v}+\vec{u}\bullet\vec{v}'$$

$$\frac{d}{dt}\left(\vec{u}\times\vec{v}\right)=\vec{u}'\times\vec{v}+\vec{u}\times\vec{v}'$$

$$\frac{d}{dt}\left(\vec{u}\left(f\left(t\right)\right)\right)=f'\left(t\right)\vec{u}'\left(f\left(t\right)\right)$$

There is also one quick definition that we should get out of the way so that we can use it when we need to.

A **smooth curve** is any curve for which $\vec{r}'\left(t\right)$ is continuous and $\vec{r}'\left(t\right)\neq 0$ for any *t* except possibly at the endpoints. A helix is a smooth curve, for example.

Finally, we need to discuss integrals of vector functions. Using both limits and derivatives as a guide it shouldn't be too surprising that we also have the following for integration for indefinite integrals

$$\int\vec{r}\left(t\right)=\left\langle\int f\left(t\right)dt,\int g\left(t\right)dt,\int h\left(t\right)dt\right\rangle+\vec{c}$$

$$\int\vec{r}\left(t\right)=\int f\left(t\right)dt\;\vec{i}+\int g\left(t\right)dt\;\vec{j}+\int h\left(t\right)dt\;\vec{k}+\vec{c}$$

and the following for definite integrals.

$$\int_a^b\vec{r}\left(t\right)dt=\left\langle\int_a^b f\left(t\right)dt,\int_a^b g\left(t\right)dt,\int_a^b h\left(t\right)dt\right\rangle$$

$$\int_a^b\vec{r}\left(t\right)dt=\int_a^b f\left(t\right)dt\;\vec{i}+\int_a^b g\left(t\right)dt\;\vec{j}+\int_a^b h\left(t\right)dt\;\vec{k}$$

With the indefinite integrals we put in a constant of integration to make sure that it was clear that the constant in this case needs to be a vector instead of a regular constant.

Also, for the definite integrals we will sometimes write it as follows,

$$\int_a^b\vec{r}\left(t\right)dt=\left(\left\langle\int f\left(t\right)dt,\int g\left(t\right)dt,\int h\left(t\right)dt\right\rangle\right)\Big|_a^b$$

$$\int_a^b\vec{r}\left(t\right)dt=\left(\int f\left(t\right)dt\;\vec{i}+\int g\left(t\right)dt\;\vec{j}+\int h\left(t\right)dt\;\vec{k}\right)\Big|_a^b$$

In other words, we will do the indefinite integral and then do the evaluation of the vector as a whole instead of on a component by component basis.

Example 3 Compute $\int \vec{r}\left(t\right)dt$ for $\vec{r}\left(t\right)=\left\langle \sin\left(t\right),6,4t\right\rangle$.

Solution

All we need to do is integrate each of the components and be done with it.

$$\int \vec{r}\left(t\right)dt=\left\langle -\cos\left(t\right),6t,2t^2\right\rangle+\vec{c}$$

Example 4 Compute $\int_0^1 \vec{r}\left(t\right)dt$ for $\vec{r}\left(t\right)=\left\langle \sin\left(t\right),6,4t\right\rangle$.

Solution

In this case all that we need to do is reuse the result from the previous example and then do the evaluation.

$$\begin{aligned}\int_0^1 \vec{r}\left(t\right)dt&=\left(\left\langle -\cos\left(t\right),6t,2t^2\right\rangle\right)\Big|_0^1\\&=\left\langle -\cos\left(1\right),6,2\right\rangle-\left\langle -1,0,0\right\rangle\\&=\left\langle 1-\cos\left(1\right),6,2\right\rangle\end{aligned}$$

Tangent, Normal and Binormal Vectors

In this section we want to look at an application of derivatives for vector functions. Actually, there are a couple of applications, but they all come back to needing the first one.

In the past we've used the fact that the derivative of a function was the slope of the tangent line. With vector functions we get exactly the same result, with one exception.

Given the vector function, $\vec{r}\left(t\right)$, we call $\vec{r}'\left(t\right)$ the **tangent vector** provided it exists and provided $\vec{r}'\left(t\right) \neq \vec{0}$. The tangent line to $\vec{r}\left(t\right)$ at *P* is then the line that passes through the point *P* and is parallel to the tangent vector, $\vec{r}'\left(t\right)$. Note that we really do need to require $\vec{r}'\left(t\right) \neq \vec{0}$ in order to have a tangent vector. If we had $\vec{r}'\left(t\right) = \vec{0}$ we would have a vector that had no magnitude and so couldn't give us the direction of the tangent.

Also, provided $\vec{r}'\left(t\right) \neq \vec{0}$, the **unit tangent vector** to the curve is given by,

$$\vec{T}\left(t\right) = \frac{\vec{r}'\left(t\right)}{\left\|\vec{r}'\left(t\right)\right\|}$$

While, the components of the unit tangent vector can be somewhat messy on occasion there are times when we will need to use the unit tangent vector instead of the tangent vector.

Example 1 Find the general formula for the tangent vector and unit tangent vector to the curve given by $\vec{r}\left(t\right) = t^2\,\vec{i} + 2\sin t\,\vec{j} + 2\cos t\,\vec{k}$.

Solution
First, by general formula we mean that we won't be plugging in a specific *t* and so we will be finding a formula that we can use at a later date if we'd like to find the tangent at any point on the curve. With that said there really isn't all that much to do at this point other than to do the work.

Here is the tangent vector to the curve.

$$\vec{r}'\left(t\right) = 2t\,\vec{i} + 2\cos t\,\vec{j} - 2\sin t\,\vec{k}$$

To get the unit tangent vector we need the length of the tangent vector.

$$\begin{aligned}\left\|\vec{r}'\left(t\right)\right\| &= \sqrt{4t^2 + 4\cos^2 t + 4\sin^2 t} \\ &= \sqrt{4t^2 + 4}\end{aligned}$$

The unit tangent vector is then,

$$\begin{aligned}\vec{T}\left(t\right) &= \frac{1}{\sqrt{4t^2+4}}\left(2t\,\vec{i} + 2\cos t\,\vec{j} - 2\sin t\,\vec{k}\right) \\ &= \frac{2t}{\sqrt{4t^2+4}}\vec{i} + \frac{2\cos t}{\sqrt{4t^2+4}}\vec{j} - \frac{2\sin t}{\sqrt{4t^2+4}}\vec{k}\end{aligned}$$

Example 2 Find the vector equation of the tangent line to the curve given by $\vec{r}\left(t\right)=t^2\,\vec{i}+2\sin t\,\vec{j}+2\cos t\,\vec{k}$ at $t=\frac{\pi}{3}$.

Solution
First we need the tangent vector and since this is the function we were working with in the previous example we can just reuse the tangent vector from that example and plug in $t=\frac{\pi}{3}$.

$$\vec{r}'\left(\frac{\pi}{3}\right)=\frac{2\pi}{3}\vec{i}+2\cos\left(\frac{\pi}{3}\right)\vec{j}-2\sin\left(\frac{\pi}{3}\right)\vec{k}=\frac{2\pi}{3}\vec{i}+\vec{j}-\sqrt{3}\,\vec{k}$$

We'll also need the point on the line at $t=\frac{\pi}{3}$ so,

$$\vec{r}\left(\frac{\pi}{3}\right)=\frac{\pi^2}{9}\vec{i}+\sqrt{3}\,\vec{j}+\vec{k}$$

The vector equation of the line is then,

$$\vec{r}\left(t\right)=\left\langle\frac{\pi^2}{9},\sqrt{3},1\right\rangle+t\left\langle\frac{2\pi}{3},1,-\sqrt{3}\right\rangle$$

Before moving on let's note a couple of things about the previous example. First, we could have used the unit tangent vector had we wanted to for the parallel vector. However, that would have made for a more complicated equation for the tangent line.

Second, notice that we used $\vec{r}\left(t\right)$ to represent the tangent line despite the fact that we used that as well for the function. Do not get excited about that. The $\vec{r}\left(t\right)$ here is much like *y* is with normal functions. With normal functions, *y* is the generic letter that we used to represent functions and $\vec{r}\left(t\right)$ tends to be used in the same way with vector functions.

Next we need to talk about the **unit normal** and the **binormal** vectors.

The unit normal vector is defined to be,

$$\vec{N}\left(t\right)=\frac{\vec{T}'\left(t\right)}{\left\|\vec{T}'\left(t\right)\right\|}$$

The unit normal is orthogonal (or normal, or perpendicular) to the unit tangent vector and hence to the curve as well. We've already seen normal vectors when we were dealing with Equations of Planes. They will show up with some regularity in several Calculus III topics.

The definition of the unit normal vector always seems a little mysterious when you first see it. It follows directly from the following fact.

Fact

Suppose that $\vec{r}\left(t\right)$ is a vector such that $\left\|\vec{r}\left(t\right)\right\|=c$ for all *t*. Then $\vec{r}'\left(t\right)$ is orthogonal to $\vec{r}\left(t\right)$.

To prove this fact is pretty simple. From the fact statement and the relationship between the magnitude of a vector and the dot product we have the following.

$$\vec{r}(t)\bullet\vec{r}(t)=\left\|\vec{r}(t)\right\|^2=c^2 \qquad \text{for all } t$$

Now, because this is true for all *t* we can see that,

$$\frac{d}{dt}\left(\vec{r}(t)\bullet\vec{r}(t)\right)=\frac{d}{dt}\left(c^2\right)=0$$

Also, recalling the fact from the previous section about differentiating a dot product we see that,

$$\frac{d}{dt}\left(\vec{r}(t)\bullet\vec{r}(t)\right)=\vec{r}'(t)\bullet\vec{r}(t)+\vec{r}(t)\bullet\vec{r}'(t)=2\vec{r}'(t)\bullet\vec{r}(t)$$

Or, upon putting all this together we get,

$$2\vec{r}'(t)\bullet\vec{r}(t)=0 \qquad\Rightarrow\qquad \vec{r}'(t)\bullet\vec{r}(t)=0$$

Therefore $\vec{r}'(t)$ is orthogonal to $\vec{r}(t)$.

The definition of the unit normal then falls directly from this. Because $\vec{T}(t)$ is a unit vector we know that $\left\|\vec{T}(t)\right\|=1$ for all *t* and hence by the Fact $\vec{T}'(t)$ is orthogonal to $\vec{T}(t)$. However, because $\vec{T}(t)$ is tangent to the curve, $\vec{T}'(t)$ must be orthogonal, or normal, to the curve as well and so be a normal vector for the curve. All we need to do then is divide by $\left\|\vec{T}'(t)\right\|$ to arrive at a unit normal vector.

Next, is the binormal vector. The binormal vector is defined to be,

$$\vec{B}(t)=\vec{T}(t)\times\vec{N}(t)$$

Because the binormal vector is defined to be the cross product of the unit tangent and unit normal vector we then know that the binormal vector is orthogonal to both the tangent vector and the normal vector.

Example 3 Find the normal and binormal vectors for $\vec{r}(t)=\left\langle t,3\sin t,3\cos t\right\rangle$.

Solution

We first need the unit tangent vector so first get the tangent vector and its magnitude.

$$\vec{r}'(t)=\left\langle 1,3\cos t,-3\sin t\right\rangle$$

$$\left\|\vec{r}'(t)\right\|=\sqrt{1+9\cos^2 t+9\sin^2 t}=\sqrt{10}$$

The unit tangent vector is then,

$$\vec{T}(t)=\left\langle\frac{1}{\sqrt{10}},\frac{3}{\sqrt{10}}\cos t,-\frac{3}{\sqrt{10}}\sin t\right\rangle$$

The unit normal vector will now require the derivative of the unit tangent and its magnitude.

$$\vec{T}'(t) = \left\langle 0, -\frac{3}{\sqrt{10}}\sin t, -\frac{3}{\sqrt{10}}\cos t \right\rangle$$

$$\left\| \vec{T}'(t) \right\| = \sqrt{\frac{9}{10}\sin^2 t + \frac{9}{10}\cos^2 t} = \sqrt{\frac{9}{10}} = \frac{3}{\sqrt{10}}$$

The unit normal vector is then,

$$\vec{N}(t) = \frac{\sqrt{10}}{3}\left\langle 0, -\frac{3}{\sqrt{10}}\sin t, -\frac{3}{\sqrt{10}}\cos t \right\rangle = \left\langle 0, -\sin t, -\cos t \right\rangle$$

Finally, the binormal vector is,

$$\begin{aligned}
\vec{B}(t) &= \vec{T}(t) \times \vec{N}(t) \\
&= \left| \begin{array}{ccc} \vec{i} & \vec{j} & \vec{k} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}}\cos t & -\frac{3}{\sqrt{10}}\sin t \\ 0 & -\sin t & -\cos t \end{array} \right| \begin{array}{cc} \vec{i} & \vec{j} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}}\cos t \\ 0 & -\sin t \end{array} \\
&= -\frac{3}{\sqrt{10}}\cos^2 t\,\vec{i} - \frac{1}{\sqrt{10}}\sin t\,\vec{k} + \frac{1}{\sqrt{10}}\cos t\;\vec{j} - \frac{3}{\sqrt{10}}\sin^2 t\,\vec{i} \\
&= -\frac{3}{\sqrt{10}}\vec{i} + \frac{1}{\sqrt{10}}\cos t\;\vec{j} - \frac{1}{\sqrt{10}}\sin t\;\vec{k}
\end{aligned}$$

Arc Length with Vector Functions

In this section we'll recast an old formula into terms of vector functions. We want to determine the length of a vector function,

$$\vec{r}(t) = \langle f(t), g(t), h(t) \rangle$$

on the interval $a \le t \le b$.

We actually already know how to do this. Recall that we can write the vector function into the parametric form,

$$x = f(t) \qquad y = g(t) \qquad z = h(t)$$

Also, recall that with two dimensional parametric curves the arc length is given by,

$$L = \int_a^b \sqrt{\left[f'(t)\right]^2 + \left[g'(t)\right]^2}\, dt$$

There is a natural extension of this to three dimensions. So, the length of the curve $\vec{r}(t)$ on the interval $a \le t \le b$ is,

$$L = \int_a^b \sqrt{\left[f'(t)\right]^2 + \left[g'(t)\right]^2 + \left[h'(t)\right]^2}\, dt$$

There is a nice simplification that we can make for this. Notice that the integrand (the function we're integrating) is nothing more than the magnitude of the tangent vector,

$$\left\|\vec{r}'(t)\right\| = \sqrt{\left[f'(t)\right]^2 + \left[g'(t)\right]^2 + \left[h'(t)\right]^2}$$

Therefore, the arc length can be written as,

$$L = \int_a^b \left\|\vec{r}'(t)\right\| dt$$

Let's work a quick example of this.

Example 1 Determine the length of the curve $\vec{r}(t) = \langle 2t, 3\sin(2t), 3\cos(2t) \rangle$ on the interval $0 \le t \le 2\pi$.

Solution
We will first need the tangent vector and its magnitude.

$$\vec{r}'(t) = \langle 2, 6\cos(2t), -6\sin(2t) \rangle$$

$$\left\|\vec{r}'(t)\right\| = \sqrt{4 + 36\cos^2(2t) + 36\sin^2(2t)} = \sqrt{4+36} = 2\sqrt{10}$$

The length is then,

$$\begin{aligned} L &= \int_a^b \left\| \vec{r}'(t) \right\| dt \\ &= \int_0^{2\pi} 2\sqrt{10}\, dt \\ &= 4\pi\sqrt{10} \end{aligned}$$

We need to take a quick look at another concept here. We define the **arc length function** as,

$$s(t) = \int_0^t \left\| \vec{r}'(u) \right\| du$$

Before we look at why this might be important let's work a quick example.

Example 2 Determine the arc length function for $\vec{r}(t) = \left\langle 2t, 3\sin(2t), 3\cos(2t) \right\rangle$.

Solution
From the previous example we know that,

$$\left\| \vec{r}'(t) \right\| = 2\sqrt{10}$$

The arc length function is then,

$$s(t) = \int_0^t 2\sqrt{10}\, du = \left(2\sqrt{10}\, u \right)\Big|_0^t = 2\sqrt{10}\, t$$

Okay, just why would we want to do this? Well let's take the result of the example above and solve it for *t*.

$$t = \frac{s}{2\sqrt{10}}$$

Now, taking this and plugging it into the original vector function and we can **reparameterize** the function into the form, $\vec{r}(t(s))$. For our function this is,

$$\vec{r}(t(s)) = \left\langle \frac{s}{\sqrt{10}}, 3\sin\left(\frac{s}{\sqrt{10}} \right), 3\cos\left(\frac{s}{\sqrt{10}} \right) \right\rangle$$

So, why would we want to do this? Well with the reparameterization we can now tell where we are on the curve after we've traveled a distance of *s* along the curve. Note as well that we will start the measurement of distance from where we are at $t = 0$.

Example 3 Where on the curve $\vec{r}(t) = \left\langle 2t, 3\sin(2t), 3\cos(2t) \right\rangle$ are we after traveling for a distance of $\dfrac{\pi\sqrt{10}}{3}$?

Solution
To determine this we need the reparameterization, which we have from above.

$$\vec{r}\left(t\left(s\right)\right)=\left\langle\frac{s}{\sqrt{10}},3\sin\left(\frac{s}{\sqrt{10}}\right),3\cos\left(\frac{s}{\sqrt{10}}\right)\right\rangle$$

Then, to determine where we are all that we need to do is plug in $s=\frac{\pi\sqrt{10}}{3}$ into this and we'll get our location.

$$\vec{r}\left(t\left(\frac{\pi\sqrt{10}}{3}\right)\right)=\left\langle\frac{\pi}{3},3\sin\left(\frac{\pi}{3}\right),3\cos\left(\frac{\pi}{3}\right)\right\rangle=\left\langle\frac{\pi}{3},\frac{3\sqrt{3}}{2},\frac{3}{2}\right\rangle$$

So, after traveling a distance of $\frac{\pi\sqrt{10}}{3}$ along the curve we are at the point $\left(\frac{\pi}{3},\frac{3\sqrt{3}}{2},\frac{3}{2}\right)$.

Curvature

In this section we want to briefly discuss the **curvature** of a smooth curve (recall that for a smooth curve we require $\vec{r}'(t)$ is continuous and $\vec{r}'(t) \neq 0$). The curvature measures how fast a curve is changing direction at a given point.

There are several formulas for determining the curvature for a curve. The formal definition of curvature is,

$$\kappa = \left| \frac{d\vec{T}}{ds} \right|$$

where $\vec{T}$ is the unit tangent and *s* is the arc length. Recall that we saw in a previous section how to reparameterize a curve to get it into terms of the arc length.

In general the formal definition of the curvature is not easy to use so there are two alternate formulas that we can use. Here they are.

$$\kappa = \frac{\left\| \vec{T}'(t) \right\|}{\left\| \vec{r}'(t) \right\|} \qquad\qquad \kappa = \frac{\left\| \vec{r}'(t) \times \vec{r}''(t) \right\|}{\left\| \vec{r}'(t) \right\|^3}$$

These may not be particularly easy to deal with either, but at least we don't need to reparameterize the unit tangent.

Example 1 Determine the curvature for $\vec{r}(t) = \langle t, 3\sin t, 3\cos t \rangle$.

Solution
Back in the section when we introduced the tangent vector we computed the tangent and unit tangent vectors for this function. These were,

$$\vec{r}'(t) = \langle 1, 3\cos t, -3\sin t \rangle$$

$$\vec{T}(t) = \left\langle \frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\cos t, -\frac{3}{\sqrt{10}}\sin t \right\rangle$$

The derivative of the unit tangent is,

$$\vec{T}'(t) = \left\langle 0, -\frac{3}{\sqrt{10}}\sin t, -\frac{3}{\sqrt{10}}\cos t \right\rangle$$

The magnitudes of the two vectors are,

$$\left\| \vec{r}'(t) \right\| = \sqrt{1 + 9\cos^2 t + 9\sin^2 t} = \sqrt{10}$$

$$\left\| \vec{T}'(t) \right\| = \sqrt{0 + \frac{9}{10}\sin^2 t + \frac{9}{10}\cos^2 t} = \sqrt{\frac{9}{10}} = \frac{3}{\sqrt{10}}$$

The curvature is then,

$$\kappa = \frac{\left\| \vec{T}'(t) \right\|}{\left\| \vec{r}'(t) \right\|} = \frac{3/\sqrt{10}}{\sqrt{10}} = \frac{3}{10}$$

In this case the curvature is constant. This means that the curve is changing direction at the same rate at every point along it. Recalling that this curve is a helix this result makes sense.

Example 2 Determine the curvature of $\vec{r}(t) = t^2\,\vec{i} + t\,\vec{k}$.

Solution

In this case the second form of the curvature would probably be easiest. Here are the first couple of derivatives.

$$\vec{r}'(t) = 2t\,\vec{i} + \vec{k} \qquad\qquad \vec{r}''(t) = 2\,\vec{i}$$

Next, we need the cross product.

$$\begin{aligned} \vec{r}'(t) \times \vec{r}''(t) &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2t & 0 & 1 \\ 2 & 0 & 0 \end{vmatrix} \begin{matrix} \vec{i} & \vec{j} \\ 2t & 0 \\ 2 & 0 \end{matrix} \\ &= 2\,\vec{j} \end{aligned}$$

The magnitudes are,

$$\left\| \vec{r}'(t) \times \vec{r}''(t) \right\| = 2 \qquad\qquad \left\| \vec{r}'(t) \right\| = \sqrt{4t^2 + 1}$$

The curvature at any value of t is then,

$$\kappa = \frac{2}{\left(4t^2 + 1\right)^{\frac{3}{2}}}$$

There is a special case that we can look at here as well. Suppose that we have a curve given by $y = f(x)$ and we want to find its curvature.

As we saw when we first looked at vector functions we can write this as follows,

$$\vec{r}(x) = x\,\vec{i} + f(x)\,\vec{j}$$

If we then use the second formula for the curvature we will arrive at the following formula for the curvature.

$$\kappa = \frac{\left| f''(x) \right|}{\left(1 + \left[f'(x) \right]^2\right)^{\frac{3}{2}}}$$

Velocity and Acceleration

In this section we need to take a look at the velocity and acceleration of a moving object.

From Calculus I we know that given the position function of an object that the velocity of the object is the first derivative of the position function and the acceleration of the object is the second derivative of the position function.

So, given this it shouldn't be too surprising that if the position function of an object is given by the vector function $\vec{r}\left(t\right)$ then the velocity and acceleration of the object is given by,

$$\vec{v}\left(t\right)=\vec{r}\,'\left(t\right) \qquad\qquad \vec{a}\left(t\right)=\vec{r}\,''\left(t\right)$$

Notice that the velocity and acceleration are also going to be vectors as well.

In the study of the motion of objects the acceleration is often broken up into a **tangential component**, a_T, and a **normal component**, a_N. The tangential component is the part of the acceleration that is tangential to the curve and the normal component is the part of the acceleration that is normal (or orthogonal) to the curve. If we do this we can write the acceleration as,

$$\vec{a}=a_T\vec{T}+a_N\vec{N}$$

where $\vec{T}$ and $\vec{N}$ are the unit tangent and unit normal for the position function.

If we define $v=\left\|\vec{v}\left(t\right)\right\|$ then the tangential and normal components of the acceleration are given by,

$$a_T=v'=\frac{\vec{r}\,'\left(t\right)\bullet\vec{r}\,''\left(t\right)}{\left\|\vec{r}\,'\left(t\right)\right\|} \qquad\qquad a_N=\kappa v^2=\frac{\left\|\vec{r}\,'\left(t\right)\times\vec{r}\,''\left(t\right)\right\|}{\left\|\vec{r}\,'\left(t\right)\right\|}$$

where κ is the curvature for the position function.

There are two formulas to use here for each component of the acceleration and while the second formula may seem overly complicated it is often the easier of the two. In the tangential component, v, may be messy and computing the derivative may be unpleasant. In the normal component we will already be computing both of these quantities in order to get the curvature and so the second formula in this case is definitely the easier of the two.

Let's take a quick look at a couple of examples.

Example 1 If the acceleration of an object is given by $\vec{a}=\vec{i}+2\vec{j}+6t\vec{k}$ find the object's velocity and position functions given that the initial velocity is $\vec{v}\left(0\right)=\vec{j}-\vec{k}$ and the initial position is $\vec{r}\left(0\right)=\vec{i}-2\vec{j}+3\vec{k}$.

Solution

We'll first get the velocity. To do this all (well almost all) we need to do is integrate the acceleration.

$$\begin{aligned}\vec{v}(t) &= \int \vec{a}(t)\,dt \\ &= \int \vec{i} + 2\vec{j} + 6t\vec{k}\,dt \\ &= t\,\vec{i} + 2t\,\vec{j} + 3t^2\,\vec{k} + \vec{c}\end{aligned}$$

To completely get the velocity we will need to determine the "constant" of integration. We can use the initial velocity to get this.

$$\vec{j} - \vec{k} = \vec{v}(0) = \vec{c}$$

The velocity of the object is then,

$$\begin{aligned}\vec{v}(t) &= t\,\vec{i} + 2t\,\vec{j} + 3t^2\,\vec{k} + \vec{j} - \vec{k} \\ &= t\,\vec{i} + (2t+1)\,\vec{j} + (3t^2-1)\vec{k}\end{aligned}$$

We will find the position function by integrating the velocity function.

$$\begin{aligned}\vec{r}(t) &= \int \vec{v}(t)\,dt \\ &= \int t\,\vec{i} + (2t+1)\,\vec{j} + (3t^2-1)\vec{k}\,dt \\ &= \frac{1}{2}t^2\,\vec{i} + (t^2+t)\,\vec{j} + (t^3-t)\vec{k} + \vec{c}\end{aligned}$$

Using the initial position gives us,

$$\vec{i} - 2\vec{j} + 3\vec{k} = \vec{r}(0) = \vec{c}$$

So, the position function is,

$$\vec{r}(t) = \left(\frac{1}{2}t^2 + 1\right)\vec{i} + (t^2+t-2)\,\vec{j} + (t^3-t+3)\vec{k}$$

Example 2 For the object in the previous example determine the tangential and normal components of the acceleration.

Solution

There really isn't much to do here other than plug into the formulas. To do this we'll need to notice that,

$$\begin{aligned}\vec{r}\,'(t) &= t\,\vec{i} + (2t+1)\,\vec{j} + (3t^2-1)\vec{k} \\ \vec{r}\,''(t) &= \vec{i} + 2\vec{j} + 6t\vec{k}\end{aligned}$$

Let's first compute the dot product and cross product that we'll need for the formulas.

$$\vec{r}\,'(t)\bullet\vec{r}\,''(t) = t + 2(2t+1) + 6t(3t^2-1) = 18t^3 - t + 2$$

$$\vec{r}'(t)\times\vec{r}''(t)=\left|\begin{array}{ccc}\vec{i} & \vec{j} & \vec{k}\\ t & 2t+1 & 3t^2-1\\ 1 & 2 & 6t\end{array}\right|\begin{array}{cc}\vec{i} & \vec{j}\\ t & 2t+1\\ 1 & 2\end{array}$$

$$=(6t)(2t+1)\vec{i}+\left(3t^2-1\right)\vec{j}+2t\vec{k}-6t^2\vec{j}-2\left(3t^2-1\right)\vec{i}-(2t+1)\vec{k}$$

$$=\left(6t^2+6t+2\right)\vec{i}-\left(3t^2+1\right)\vec{j}-\vec{k}$$

Next, we also need a couple of magnitudes.

$$\left\|\vec{r}'(t)\right\|=\sqrt{t^2+(2t+1)^2+\left(3t^2-1\right)^2}=\sqrt{9t^4-t^2+4t+2}$$

$$\left\|\vec{r}'(t)\times\vec{r}''(t)\right\|=\sqrt{\left(6t^2+6t+2\right)^2+\left(3t^2+1\right)^2+1}=\sqrt{45t^4+72t^3+66t^2+24t+6}$$

The tangential component of the acceleration is then,

$$a_T=\frac{18t^3-t+2}{\sqrt{9t^4-t^2+4t+2}}$$

The normal component of the acceleration is,

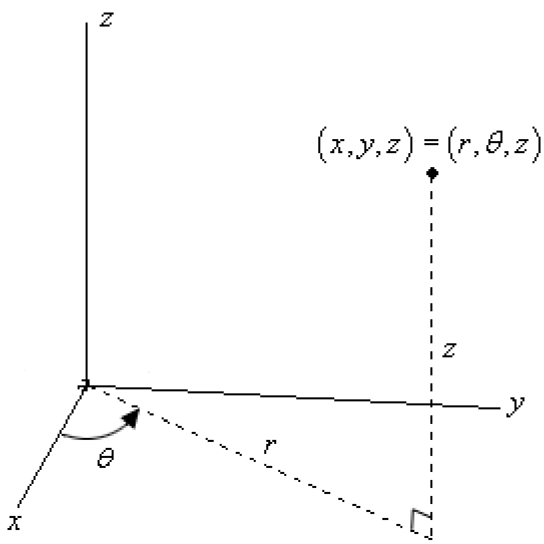
$$a_N=\frac{\sqrt{45t^4+72t^3+66t^2+24t+6}}{\sqrt{9t^4-t^2+4t+2}}=\sqrt{\frac{45t^4+72t^3+66t^2+24t+6}{9t^4-t^2+4t+2}}$$

Cylindrical Coordinates

As with two dimensional space the standard (x, y, z) coordinate system is called the Cartesian coordinate system. In the last two sections of this chapter we'll be looking at some alternate coordinate systems for three dimensional space.

We'll start off with the cylindrical coordinate system. This one is fairly simple as it is nothing more than an extension of polar coordinates into three dimensions. Not only is it an extension of polar coordinates, but we extend it into the third dimension just as we extend Cartesian coordinates into the third dimension. All that we do is add a z on as the third coordinate. The r and θ are the same as with polar coordinates.

Here is a sketch of a point in $\mathbb{R}^3$.



The conversions for x and y are the same conversions that we used back when we were looking at polar coordinates. So, if we have a point in cylindrical coordinates the Cartesian coordinates can be found by using the following conversions.

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}$$

The third equation is just an acknowledgement that the z-coordinate of a point in Cartesian and polar coordinates is the same.

Likewise, if we have a point in Cartesian coordinates the cylindrical coordinates can be found by using the following conversions.

$$r = \sqrt{x^2 + y^2} \qquad \text{OR} \qquad r^2 = x^2 + y^2$$
$$\theta = \tan^{-1}\left(\frac{y}{x}\right)$$
$$z = z$$

Let's take a quick look at some surfaces in cylindrical coordinates.

Example 1 Identify the surface for each of the following equations.

(a) $r = 5$
(b) $r^2 + z^2 = 100$
(c) $z = r$

Solution

(a) In two dimensions we know that this is a circle of radius 5. Since we are now in three dimensions and there is no *z* in equation this means it is allowed to vary freely. So, for any given *z* we will have a circle of radius 5 centered on the *z*-axis.

In other words, we will have a cylinder of radius 5 centered on the *z*-axis.

(b) This equation will be easy to identify once we convert back to Cartesian coordinates.

$$r^2 + z^2 = 100$$
$$x^2 + y^2 + z^2 = 100$$

So, this is a sphere centered at the origin with radius 10.

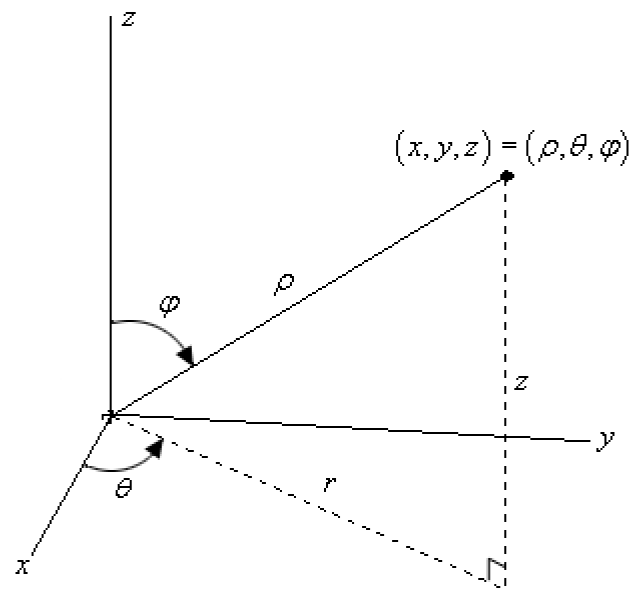
(c) Again, this one won't be too bad if we convert back to Cartesian. For reasons that will be apparent eventually, we'll first square both sides, then convert.

$$z^2 = r^2$$
$$z^2 = x^2 + y^2$$

From the section on quadric surfaces we know that this is the equation of a cone.

Spherical Coordinates

In this section we will introduce spherical coordinates. Spherical coordinates can take a little getting used to. It's probably easiest to start things off with a sketch.



Spherical coordinates consist of the following three quantities.

First there is ρ. This is the distance from the origin to the point and we will require $\rho \geq 0$.

Next there is θ. This is the same angle that we saw in polar/cylindrical coordinates. It is the angle between the positive *x*-axis and the line above denoted by *r* (which is also the same *r* as in polar/cylindrical coordinates). There are no restrictions on θ.

Finally there is φ. This is the angle between the positive *z*-axis and the line from the origin to the point. We will require $0 \leq \varphi \leq \pi$.

In summary, ρ is the distance from the origin to the point, φ is the angle that we need to rotate down from the positive *z*-axis to get to the point and θ is how much we need to rotate around the *z*-axis to get to the point.

We should first derive some conversion formulas. Let's first start with a point in spherical coordinates and ask what the cylindrical coordinates of the point are. So, we know (ρ, θ, φ) and want to find (r, θ, z). Of course we really only need to find *r* and *z* since θ is the same in both coordinate systems.

We will be able to do all of our work by looking at the right triangle shown above in our sketch. With a little geometry we see that the angle between *z* and ρ is φ and so we can see that,

$$z = \rho\cos\varphi$$
$$r = \rho\sin\varphi$$

and these are exactly the formulas that we were looking for. So, given a point in spherical coordinates the cylindrical coordinates of the point will be,

$$\begin{aligned} r &= \rho\sin\varphi \\ \theta &= \theta \\ z &= \rho\cos\varphi \end{aligned}$$

Note as well that,

$$r^2 + z^2 = \rho^2\cos^2\varphi + \rho^2\sin^2\varphi = \rho^2\left(\cos^2\varphi + \sin^2\varphi\right) = \rho^2$$

Or,

$$\rho^2 = r^2 + z^2$$

Next, let's find the Cartesian coordinates of the same point. To do this we'll start with the cylindrical conversion formulas from the previous section.

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}$$

Now all that we need to do is use the formulas from above for r and z to get,

$$\begin{aligned} x &= \rho\sin\varphi\cos\theta \\ y &= \rho\sin\varphi\sin\theta \\ z &= \rho\cos\varphi \end{aligned}$$

Also note that since we know that $r^2 = x^2 + y^2$ we get,

$$\rho^2 = x^2 + y^2 + z^2$$

Converting points from Cartesian or cylindrical coordinates into spherical coordinates is usually done with the same conversion formulas. To see how this is done let's work an example of each.

Example 1 Perform each of the following conversions.

(a) Convert the point $\left(\sqrt{6}, \frac{\pi}{4}, \sqrt{2}\right)$ from cylindrical to spherical coordinates.

[Solution]

(b) Convert the point $\left(-1, 1, -\sqrt{2}\right)$ from Cartesian to spherical coordinates.

[Solution]

Solution

(a) Convert the point $\left(\sqrt{6}, \frac{\pi}{4}, \sqrt{2}\right)$ **from cylindrical to spherical coordinates.**

We'll start by acknowledging that θ is the same in both coordinate systems and so we don't need to do anything with that.

Next, let's find ρ.

$$\rho = \sqrt{r^2 + z^2} = \sqrt{6+2} = \sqrt{8} = 2\sqrt{2}$$

Finally, let's get φ. To do this we can use either the conversion for *r* or *z*. We'll use the conversion for *z*.

$$z = \rho\cos\varphi \qquad \Rightarrow \qquad \cos\varphi = \frac{z}{\rho} = \frac{\sqrt{2}}{2\sqrt{2}} \qquad \Rightarrow \qquad \varphi = \cos^{-1}\left(\frac{1}{2}\right) = \frac{\pi}{3}$$

Notice that there are many possible values of φ that will give $\cos\varphi = \frac{1}{2}$, however, we have restricted φ to the range $0 \le \varphi \le \pi$ and so this is the only possible value in that range.

So, the spherical coordinates of this point will are $\left(2\sqrt{2}, \frac{\pi}{4}, \frac{\pi}{3}\right)$.

[Return to Problems]

(b) Convert the point $\left(-1, 1, -\sqrt{2}\right)$ **from Cartesian to spherical coordinates.**

The first thing that we'll do here is find ρ.

$$\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{1+1+2} = 2$$

Now we'll need to find φ. We can do this using the conversion for *z*.

$$z = \rho\cos\varphi \qquad \Rightarrow \qquad \cos\varphi = \frac{z}{\rho} = \frac{-\sqrt{2}}{2} \qquad \Rightarrow \qquad \varphi = \cos^{-1}\left(\frac{-\sqrt{2}}{2}\right) = \frac{3\pi}{4}$$

As with the last parts this will be the only possible φ in the range allowed.

Finally, let's find θ. To do this we can use the conversion for *x* or *y*. We will use the conversion for *y* in this case.

$$\sin\theta = \frac{y}{\rho\sin\varphi} = \frac{1}{2\left(\frac{\sqrt{2}}{2}\right)} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \qquad \Rightarrow \qquad \theta = \frac{\pi}{4} \text{ or } \theta = \frac{3\pi}{4}$$

Now, we actually have more possible choices for θ but all of them will reduce down to one of the two angles above since they will just be one of these two angles with one or more complete rotations around the unit circle added on.

We will however, need to decide which one is the correct angle since only one will be. To do this let's notice that, in two dimensions, the point with coordinates $x=-1$ and $y=1$ lies in the second quadrant. This means that θ must be angle that will put the point into the second quadrant. Therefore, the second angle, $\theta=\frac{3\pi}{4}$, must be the correct one.

The spherical coordinates of this point are then $\left(2,\frac{3\pi}{4},\frac{3\pi}{4}\right)$.

[Return to Problems]

Now, let's take a look at some equations and identify the surfaces that they represent.

Example 2 Identify the surface for each of the following equations.

(a) $\rho=5$ [Solution]

(b) $\varphi=\dfrac{\pi}{3}$ [Solution]

(c) $\theta=\dfrac{2\pi}{3}$ [Solution]

(d) $\rho\sin\varphi=2$ [Solution]

Solution

(a) $\rho=5$

There are a couple of ways to think about this one.

First, think about what this equation is saying. This equation says that, no matter what θ and φ are, the distance from the origin must be 5. So, we can rotate as much as we want away from the *z*-axis and around the *z*-axis, but we must always remain at a fixed distance from the origin. This is exactly what a sphere is. So, this is a sphere of radius 5 centered at the origin.

The other way to think about it is to just convert to Cartesian coordinates.

$$\begin{aligned}\rho&=5\\ \rho^2&=25\\ x^2+y^2+z^2&=25\end{aligned}$$

Sure enough a sphere of radius 5 centered at the origin.

[Return to Problems]

(b) $\varphi=\dfrac{\pi}{3}$

In this case there isn't an easy way to convert to Cartesian coordinates so we'll just need to think about this one a little. This equation says that no matter how far away from the origin that we move and no matter how much we rotate around the *z*-axis the point must always be at an angle of $\frac{\pi}{3}$ from the *z*-axis.

This is exactly what happens in a cone. All of the points on a cone are a fixed angle from the *z*-

axis. So, we have a cone whose points are all at an angle of $\frac{\pi}{3}$ from the z-axis.

[Return to Problems]

(c) $\theta = \dfrac{2\pi}{3}$

As with the last part we won't be able to easily convert to Cartesian coordinates here. In this case no matter how far from the origin we get or how much we rotate down from the positive z-axis the points must always form an angle of $\frac{2\pi}{3}$ with the x-axis.

Points in a vertical plane will do this. So, we have a vertical plane that forms an angle of $\frac{2\pi}{3}$ with the positive x-axis.

[Return to Problems]

(d) $\rho \sin\varphi = 2$

In this case we can convert to Cartesian coordinates so let's do that. There are actually two ways to do this conversion. We will look at both since both will be used on occasion.

Solution 1
In this solution method we will convert directly to Cartesian coordinates. To do this we will first need to square both sides of the equation.

$$\rho^2 \sin^2\varphi = 4$$

Now, for no apparent reason add $\rho^2\cos^2\varphi$ to both sides.

$$\begin{aligned} \rho^2 \sin^2\varphi + \rho^2\cos^2\varphi &= 4 + \rho^2\cos^2\varphi \\ \rho^2\left(\sin^2\varphi + \cos^2\varphi\right) &= 4 + \rho^2\cos^2\varphi \\ \rho^2 &= 4 + \left(\rho\cos\varphi\right)^2 \end{aligned}$$

Now we can convert to Cartesian coordinates.

$$\begin{aligned} x^2 + y^2 + z^2 &= 4 + z^2 \\ x^2 + y^2 &= 4 \end{aligned}$$

So, we have a cylinder of radius 2 centered on the z-axis.

This solution method wasn't too bad, but it did require some not so obvious steps to complete.

Solution 2
This method is much shorter, but also involves something that you may not see the first time around. In this case instead of going straight to Cartesian coordinates we'll first convert to cylindrical coordinates.

This won't always work, but in this case all we need to do is recognize that $r = \rho\sin\varphi$ and we will get something we can recognize. Using this we get,

$$\rho \sin\varphi = 2$$
$$r = 2$$

At this point we know this is a cylinder (remember that we're in three dimensions and so this isn't a circle!). However, let's go ahead and finish the conversion process out.

$$r^2 = 4$$
$$x^2 + y^2 = 4$$

[Return to Problems]

So, as we saw in the last part of the previous example it will *sometimes* be easier to convert equations in spherical coordinates into cylindrical coordinates before converting into Cartesian coordinates. This won't always be easier, but it can make some of the conversions quicker and easier.

The last thing that we want to do in this section is generalize the first three parts of the previous example.

$\rho = a$	sphere of radius a centered at the origin
$\varphi = \alpha$	cone that makes an angle of α with the positive $z-$axis
$\theta = \beta$	vertical plane that makes an angle of β with the positive $x-$axis