# Math 119

# Calculus II for Engineering

*Lecture Notes*

David Harmsworth

Department of Applied Mathematics

University of Waterloo

2014, 2018

## Part I

# Multivariate Calculus

## 1 Introduction

In Math 117 we discussed the essentials of the calculus of functions of one variable, and relationships between variables defined by those functions: $y = f(x)$, or

$$x \to \boxed{f} \to y.$$

We could also write $f : \mathbb{R} \to \mathbb{R}$, of course, to denote that $f$ takes one (real) input and gives one (real) output. Although, as we've seen, there are lots of things we can do with these functions, many real-world phenomena depend on multiple variables, and so in the first half of Math 119, we will turn our attention to functions which take multiple variables as input. As a couple of simple motivating examples, consider that the altitude of the earth's surface could be viewed as a function of both latitude and longitude, while the temperature in a region of space could be considered to be a function of four variables (three for the spatial dimensions, and one for time).

Most of our examples in this course will feature functions of just two variables. We'll follow the traditional notation and use $x$ and $y$ for the two input variables, and $z$ for the output:

$$\left.\begin{matrix} x \\ y \end{matrix}\right\} \to \boxed{f} \to z.$$

We'll usually write $z = f(x, y)$, and to indicate that a function $f$ is of this type without specifying names for the variables we might write $f : \mathbb{R}^2 \to \mathbb{R}$. Most of the definitions and theorems we'll discuss can be generalized fairly easily to functions of *more* than two variables, and we will in fact consider the three-variable case on occasion, writing perhaps $w = f(x, y, z)$ or $f : \mathbb{R}^3 \to \mathbb{R}$.

Observe that while the *range* of a multivariate function is always a subset of $\mathbb{R}$, the *domain* will be a subset of $\mathbb{R}^n$. This means that our graphs will no longer be curves over intervals, but surfaces and "hypersurfaces" over regions! Furthermore, in developing the theory of multivari-

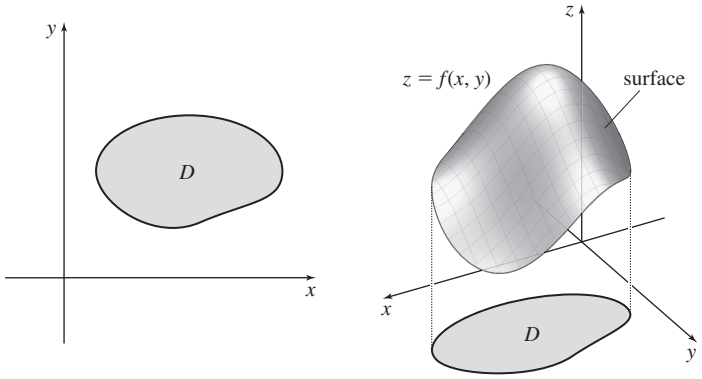
ate calculus, we will often find it useful to view the input as a *vector*; at this point calculus and linear algebra begin to intertwine.

**Aside:** Strictly speaking, this kind of function should be referred to as a *scalar field*. There are a couple of other types of "functions" in multivariable calculus as well, and when we're trying to avoid confusion these are the terms we'll use:

| | |
|---|---|
| (scalar) function: $\mathbb{R} \to \mathbb{R}$ | vector function: $\mathbb{R} \to \mathbb{R}^n$ |
| scalar field: $\mathbb{R}^n \to \mathbb{R}$ | vector field: $\mathbb{R}^n \to \mathbb{R}^n$ |

**Graphs of Scalar Fields**

For a two-variable scalar field $z = f(x, y)$, the domain is a region in the $xy$-plane. To graph this function, we need a third co-ordinate axis, so the graph is a surface in $\mathbb{R}^3$:

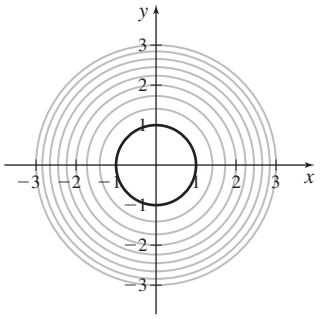



For a function of three variables, say $w = f(x, y, z)$, the graph would be a "hypersurface" existing in four dimensions; there are probably very few people who can visualize these effectively! Even three-dimensional structures can be difficult to represent clearly in two dimensions, so we often use *contour plots* instead. This idea should be a familiar one:

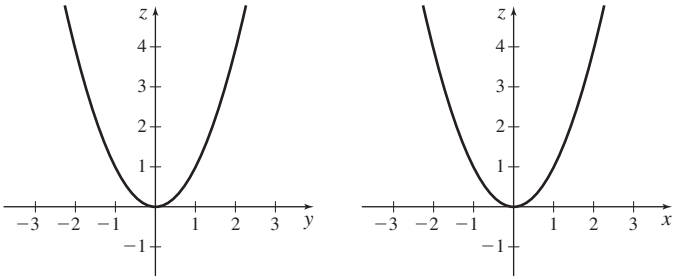
- A topographical map is a contour plot of altitude as a function of latitude and longitude
- The isobars / isotherms on weather maps are contour plots of air pressure / air temperature (again as functions of latitude and longitude).

To generate a contour plot, we simply use $f(x, y) = K$. For each value we assign to $K$, we obtain the equation of a *level curve* of $f$ (in $\mathbb{R}^2$).

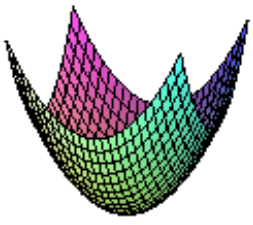
**Example:** Consider the field $f(x, y) = x^2 + y^2$ (or the relation $z = x^2 + y^2$). Setting $f = K$ gives us the expression $x^2 + y^2 = K$, so the level curves are all circles of radius $\sqrt{K}$:



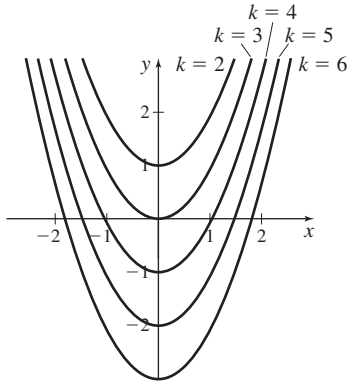

This will often be enough for us to visualize the surface (we should label the $z$-values corresponding to each curve, but we haven't done so in this figure). If it's still not clear, we can also consider cross-sections orthogonal to the other two axes. One value for each of $x$ and $y$ is usually enough, so we would traditionally set $x = 0$, which gives the equation $z = y^2$, and $y = 0$, which gives the similar equation $z = x^2$:



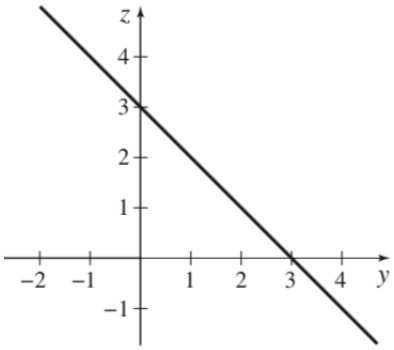

From these two figures you should be able to construct a mental image of the surface; you can see the full graph below (this plot doesn't show the contour lines, and was generated using a square domain):

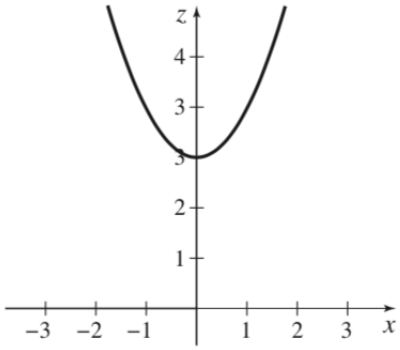


**Example:** Consider the function $f(x, y) = x^2 - y + 3$. This time setting $f = K$ gives $x^2 - y + 3 = K$, or $y = x^2 + 3 - K$, so the level curves form a family of parabolas (and this time we've labelled them):




If that's not enough to give you the full picture, you can consider the other cross-sections: setting $x = 0$ gives $z = 3 - y$, while setting $y = 0$ gives $z = x^2 + 3$:

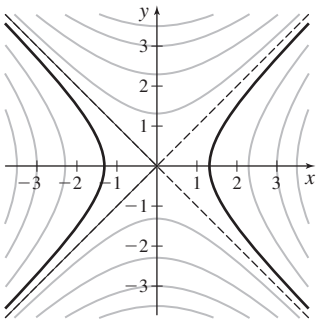



That *should* be enough to allow you to picture the surface. (Hint: the $y = 0$ cross-section might actually throw you off here; concentrate on the level curves and the straight line given by $x = 0$.)

**Example:** This idea can help us to understand the behaviour of functions of three variables as well. In this case setting $f = K$ produces *level surfaces*. For example, for the equation $w = x^2 + y^2 + z^2$ we find the level surfaces to be the concentric spheres $x^2 + y^2 + z^2 = K$. Even though the graph of this function has four dimensions, so we might not be able to truly visualize it, we have a good idea of how it behaves from its level surfaces.

**Example:** Finally, consider the function $f(x, y) = x^2 - y^2$. The level curves in this case have equation $x^2 - y^2 = K$; you should recognize these as being hyperbolas. If $K > 0$ they open leftwards and rightwards, while if $K < 0$ they open upwards and downwards. In the special case where $K = 0$ we find the pair of straight lines $y = \pm x$:




If this image confuses you, investigate the other cross-sections, and you should be able to identify the shape of the graph of $f$.

# 2 Limits and Partial Derivatives

## 2.1 Multivariate Limits (optional section)

Scalar fields can exhibit some perplexing behaviour near discontinuities. Consider the following examples:

**Example:** Does the function $f(x,y) = \dfrac{2xy}{x^2+y^2}$ have a limit as the input point $(x,y)$ approaches the origin? If you think about this for a moment, you'll start to appreciate the complexity of this sort of question. Are we supposed to keep $y$ fixed and let $x$ approach zero? Keep $x$ fixed and let $y$ approach zero? Let both approach zero simultaneously? Well, if the limit exists, it shouldn't matter how we approach the origin; we should always get the same result. So, let's consider a couple of different ways to approach the origin:

- If we fix the value of $y$ as zero, then $f(x,y) = 0$ for all $(x,y) \neq (0,0)$, and therefore the limit is zero.
- If we fix the value of $x$ as zero, then we get the same result. Can we conclude then that the limit is zero? Not yet...
- If we set $y = x$, then $f(x,y) = f(x,x) = \dfrac{2x^2}{2x^2} = 1$ for all $(x,y) \neq (0,0)$, and so the limit along this route to the origin is *one*.

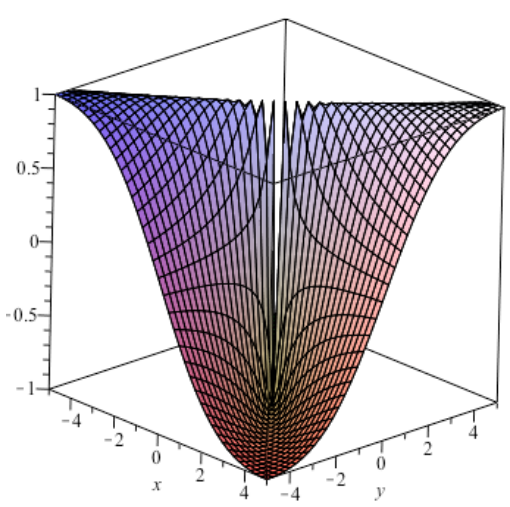
Since we obtain different limits along different paths toward the origin, we conclude that the limit does not exist.

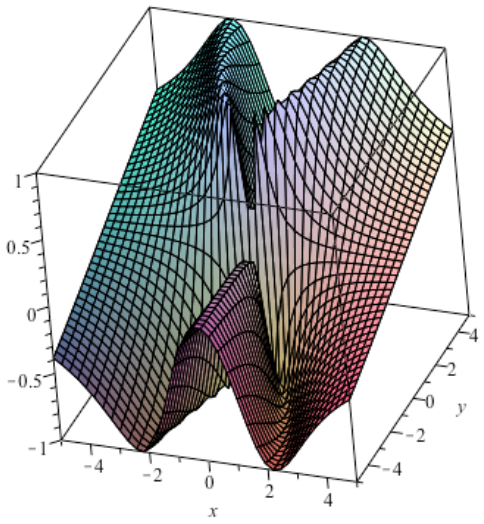
**Example:** Now consider the function $f(x,y) = \dfrac{2x^2y}{x^4+y^2}$. This time

- If we set $y = 0$ and let $x \to 0$ we find $f \to 0$.
- If we set $x = 0$ and let $y \to 0$ we find $f \to 0$.
- If we set $y = x$ then $f(x,y) = f(x,x) = \dfrac{2x^3}{x^4+x^2} = \dfrac{2x}{x^2+1}$, and as $x \to 0$ we find again that $f \to 0$.
- In fact, we can set $y = kx$, giving $f(x,y) = f(x,x) = \dfrac{2kx^3}{x^4+k^2x^2} = \dfrac{2kx}{x^2+k^2}$; this tells us that we can approach the origin *along* ***any*** *straight line*, and the limit will be zero.

It sounds as if that should do it. However, there *are* other ways to approach the origin! Suppose we approach along the parabola with equation $y = x^2$. Along that parabola our output is $f(x, x^2) = \dfrac{2x^4}{x^4 + x^4} = 1$, so the limit along this route to the origin is *one*! So, approaching along this parabola gives a different limit than approaching along a line, and we must therefore conclude that $\lim\limits_{(x,y)\to(0,0)} \dfrac{2x^2y}{x^4 + y^2}$ *does not exist.*

These surfaces can be plotted (see below) but it isn't always easy to comprehend this sort of behaviour!




So, if a limit *does* exist, how do we prove it? Somehow we have to show that the limit will be the same along every conceivable route toward the point in question. For that, the only really effective tool we have available is the Squeeze Theorem.

Fortunately, you will rarely be confronted with limit problems of this type. As you'll see, we can extend our concepts of derivatives and integrals to multivariate functions by using *single-variable* limits, so you can look upon these multivariate limits as a mathematical curiosity, and we can move on.

## 2.2 Partial Derivatives

How do we differentiate a function of two variables? Easily! We just "pretend" that one of the variables is a constant, and differentiate the resulting single-variable function in the usual way! (so there are two possibilities).

Technically, we have the following:

**Definition:**

The *partial derivative of* $f(x, y)$ *with respect to* $x$ *at the point* $(a, b)$ is

$$f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h},$$

if this limit exists. Similarly, the *partial derivative of* $f(x, y)$ *with respect to* $y$ *at the point* $(a, b)$ is

$$f_y(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h},$$

providing that this limit exists.

**Notation:** The notation above, $f_x$ and $f_y$, can be thought of as corresponding to the "prime" notation for ordinary derivative ($f\,'$). In the Leibniz notation $\left(\text{corresponding to } \frac{df}{dx}\right)$, we write $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ (the symbol $\partial$ is pronounced as "die"[1]). You may occasionally also see the two partial derivatives represented as $D_1$ and $D_2$.

**Note:** The definitions can be extended easily to functions of *more* than two variables. In practice, we simply "pretend" that all but one of the variables are constant, and then ordinary differentiation yields the partial derivative with respect to the remaining variable.

**Example:** Consider the function $f(x, y) = e^{xy} + \frac{x}{y}$. To find $\frac{\partial f}{\partial x}$, we view $y$ as a constant:

$$\frac{\partial f}{\partial x} = e^{xy}(y) + \frac{1}{y} = ye^{xy} + \frac{1}{y}.$$

[1]The partial derivative symbol $\partial$ was invented specifically for this purpose. It's obviously similar to the Greek letter $\delta$, the Latin letter $d$, and the Old English / Icelandic letter ð ("eth").

For $\dfrac{\partial f}{\partial y}$, we view $x$ as a constant, and differentiate the resulting function of $y$ instead:

$$\frac{\partial f}{\partial y} = e^{xy}(x) + x\left(\frac{-1}{y^2}\right) = xe^{xy} - \frac{x}{y^2}.$$

It really is that easy! (It just takes a bit of practice to get used to treating variables as constants).

**Example:** Suppose we wish to know $f_y(0,0)$, if $f(x,y) = (x^3+y^3)^{1/3}$. We can calculate the general formula for $f_y$:

$$f_y = \frac{1}{3}(x^3+y^3)^{-2/3}\cdot 3y^2 = \frac{y^2}{(x^3+y^3)^{2/3}}.$$

However, setting $x=0$, $y=0$ gives us the indeterminate form $\frac{0}{0}$. At this point it's tempting to consider the limit of $f_y$ as $(x,y)\to(0,0)$, but as we've just discussed, that is *not* an easy task (and in fact the limit doesn't exist). Instead, we can turn to the definition of $f_y$, which only requires a single-variable limit:

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

**Higher-Order Partial Derivatives**

Since $f_x(x,y)$ and $f_y(x,y)$ are each themselves functions of $x$ and $y$, they each have two partial derivatives of their own, so we expect four second-order derivatives:

**Example:** If $f(x,y) = x\cos\ y + 2x^3y$, we find

$$\frac{\partial f}{\partial x} = f_x = \cos\ y + 6x^2y,$$

and

$$\frac{\partial f}{\partial y} = f_y = -x\sin\ + 2x^3,$$

so the second-order derivatives are

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx} = 12xy$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = f_{xy} = -\sin\ y + 6x^2$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy} = -x\cos\ y$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = f_{yx} = -\sin\ y + 6x^2.$$

You'll notice that the two "mixed" partial derivatives are the same. In fact this is not a coincidence; *Clairaut's Theorem* states that if $f_x$, $f_y$, and $f_{xy}$ exist near $(a, b)$, and if $f_{xy}$ is continuous at $(a, b)$, then $f_{yx}(a, b)$ also exists, and in fact $f_{yx}(a, b) = f_{xy}(a, b)$.

In general it's safe to say that unless $f(x, y)$ is piecewise defined, and we're dealing with points on the boundary between its "pieces", the order of differentiation is unimportant.

**Example:** Suppose we wish to know $f_{yyyxx}$ for the function $f(x, y) = \ln\ |y| - xye^{y^2}$. The derivatives with respect to $y$ are complicated, but it doesn't matter! The derivatives won't exist for $y = 0$, since the original function isn't defined there anyway, but at every other point they will be continuous, and we can immediately see that *any* derivative involving two derivatives with respect to $x$ will be zero.

## 3 Tangent Planes, The Linear Approximation, and Differentials

Recall from Math 117 that the tangent line to a curve $y = f(x)$, at a point $(x_0, f(x_0))$, has equation $y = f(x_0) + f'(x_0)(x - x_0)$, and for values of $x$ near $x_0$, we can use this function as an approximation to $f$; the linear approximation is

$$L_{x_0} = f(x_0) + f'(x_0)(x - x_0).$$

We now wish to generalize this for multivariate functions.

First, realize that if a surface $z = f(x, y)$ is smooth, then at each point $(x_0, y_0)$ there should be a tangent *plane* rather than a tangent line. So, the first thing we have to do is determine what the equations of planes look like.

## Equations of Planes

Consider a non-vertical plane passing through a point $\vec{r_0} = (x_0, y_0, z_0)$ (we'll ignore vertical planes here because they can't be described using $z$ as a function of $x$ and $y$; they are vertical extensions of lines in the $xy$-plane, and so their equations will have the form $ax + by = 1$, or possibly just $x = a$ or $y = b$). We'll consider two vectors:

- Let $\vec{r} = (x, y, z)$ be another point in the plane, and consider the vector $\vec{r} - \vec{r_0}$; it will lie in the plane as well.
- Let $\hat{n} = (a, b, c)$ be a normal vector to the plane at $\vec{r_0}$.

These two vectors must be orthogonal, so $\hat{n} \cdot (\vec{r} - \vec{r_0}) = 0$. That is,

$$a\,(x - x_0) + b\,(y - y_0) + c\,(z - z_0) = 0.$$

We can rewrite this. Solving for $z$, dividing by $c$, and renaming the constants ($-\frac{a}{c} = A$,$-\frac{b}{c} = B$), we obtain

$$z = z_0 + A\,(x - x_0) + B\,(y - y_0)\,. \tag{1}$$

Alternatively, dispensing with even more of the constants, we could write $ax + by + cz = K$ (where $K$ is $ax_0 + by_0 + cz_0$).

These expressions should be reassuring; you can see that if you set $z = 0$ you obtain the equation of a line in the $xy$-plane, and in fact, if you fix the value of any one of the three variables, you obtain the equation of a line in some vertical or horizontal plane.

## Equations of Tangent Planes

Now, consider equation 1, and suppose that this is the tangent plane to a surface with equation $z = f\,(x, y)$. If we set $y = y_0$, we obtain the equation of a line in the plane $y = y_0$: $z = z_0 + A\,(x - x_0)$. From the way we've defined partial derivatives, we know that the slope of this line is $f_x\,(x_0, y_0)$, so that's the value of $A$! Similarly, if we set $x = x_0$ instead, we see that $B = f_y\,(x_0, y_0)$. Therefore, our tangent plane has equation

$$z = z_0 + f_x\,(x_0, y_0)\,(x - x_0) + f_y\,(x_0, y_0)\,(y - y_0)\,.$$

**Example:** Find the equation of the tangent plane to the surface $z = x^2y + xy^3$ at the point $(1, 2)$.

**Solution:** We simply calculate $\dfrac{\partial z}{\partial x} = 2xy + y^3$ and $\dfrac{\partial z}{\partial y} = x^2 + 3xy^2$ and evaluate them (and $f$) at the point: $f(1,2) = 10$, $f_x(1,2) = 12$, and $f_y(1,2) = 13$, so the tangent plane has equation

$$z = 10 + 12(x-1) + 13(y-2).$$

## Approximations

Now that we have the formula, using it is easy. As long as we stay close enough to the point of tangency $(x_0, y_0)$, we can say that

$$f(x,y) \approx f(x_0,y_0) + f_x(x_0,y_0) \cdot (x - x_0) + f_y(x_0,y_0) \cdot (y - y_0).$$

That's our linear approximation! For example, for the function $f(x,y) = x^2y + xy^3$ used above, we might claim that $f(1.1, 1.8) \approx 10 + 12(0.1) + 13(-0.2) = 8.6$ (in fact the exact value is 8.5932).

We can certainly use the linear approximation in this form, but just as in the single-variable case, it's often expressed in the notation of differentials. If we use $\Delta x = x - x_0$, $\Delta y = y - y_0$, and set $\Delta f = f(x,y) - f(x_0,y_0)$, then our linear approximation can be written as

$$\Delta f \;\approx\; f_x(P_0)\Delta x + f_y(P_0)\Delta y. \tag{2}$$

In the limit as $\Delta x$ and $\Delta y$ approach zero, this approximation becomes more and more accurate (the error approaches zero). To reflect this, we replace $\Delta$ with $d$, throughout, and replace the $\approx$ symbol with the equal sign. For conciseness, we also drop the "$P_0$", although it's important to remember that this is only a notational change; *the derivatives are still to be evaluated at the center of the approximation*:

$$df \;=\; f_x dx + f_y dy. \tag{3}$$

This expression is referred to as *the total differential of $f$*. It means exactly the same thing as Equation 2, except that it incorporates the information that the approximation improves as we approach the center. In practice, we often use $\Delta x$ and $\Delta y$ interchangeably with $dx$ and

$dy$, from which it follows that the actual difference $\Delta f$ can be approximated by calculating $df$. This is exactly the same practice we introduced for single variable calculus in Math 117.

**Example:** For the function $f(x, y) = e^{x+2y}$, estimate the value of $f(0.1, 0.02)$.

**Solution:** We know that $f(0, 0) = 1$, so we expect $f(0.1, 0.02)$ to be close to 1. The error in this estimate is $\Delta f$, which we can approximate using $df$. Now, $f_x(x, y) = e^{x+2y}$, and $f_y(x, y) = 2e^{x+2y}$, so $f_x(0, 0) = 1$ and $f_y(0, 0) = 2$. Therefore,

$$\begin{aligned} \Delta f \approx df &= f_x dx + f_y dy \\ &= f_x \Delta x + f_y \Delta y \\ &= 1 \cdot 0.1 + 2 \cdot 0.02 \\ &= 0.14 \end{aligned}$$

We can now conclude[2] that $f(0.1, 0.02) = f(0, 0) + \Delta f = 1 + \Delta f \approx 1 + df = 1.14$.

If you prefer, you could approach this calculation in a different way. The equation of the plane which is tangent to $f(x, y)$ at $(0, 0)$ is $z = 1+x+2y$, and we're simply using this tangent plane as an approximation to the graph of the original function. On the tangent plane, $x = 0.1$ and $y = 0.02$ give $z = 1.14$. However, the notation of differentials has advantages for problems of the type in our next example.

**Example:** The pressure $P$, volume $V$, and temperature $T$ of one mole of ideal gas obey the ideal gas law: $PV = RT$ (where $R$ is a constant). If the volume increases by 2.5% and the temperature decreases by 1.6%, what is the resulting change in pressure?

**Solution:** First, note that solving for $P$ gives $P = \dfrac{RT}{V}$. This allows us to calculate the differential as

$$\begin{aligned} dP &= \frac{\partial P}{\partial T} dT + \frac{\partial P}{\partial V} dV \\ &= \frac{R}{V} dT - \frac{RT}{V^2} dV \end{aligned}$$

[2] This is, admittedly, a bit of silly example. We're really being asked to approximate the value of $e^{0.14}$, and once we realize that it becomes a single-variable problem!

Since we're looking for the *relative* change in pressure, let's divide through by $P$ (which is equal to $\frac{RT}{V}$, of course):

$$\frac{dP}{P} = \frac{dT}{T} - \frac{dV}{V}.$$

This has conveniently given us a result in terms of exactly the information we have been given: $\frac{dP}{P} = -0.016 - 0.025 = -0.041$. That is, the pressure should decrease by about 4.1% (remember: since all of this is based upon the linear approximation, the result is itself only an *approximation*).[3]

**Example:** Suppose a company makes right-circular storage tanks that are 5 meters high with a radius of 1 meter.

**a)** How sensitive is the tank's volume to small variations in its height and radius?

**Solution:** The volume is given by $V(r,h) = \pi r^2 h$. Since the differential is $dV = \frac{\partial V}{\partial r}dr + \frac{\partial V}{\partial h}dh$, the change in volume is approximately

$$\begin{aligned} dV &= 2\pi r h dr + \pi r^2 dh \\ &= 10\pi dr + \pi dh \end{aligned}$$

Therefore the volume is 10 times more sensitive to changes in radius than to changes in height.

**b)** Now suppose that $r$ can be controlled with an error of no more than 2% (positive or negative), and $h$ to within 0.5%. What is the greatest possible percentage error in the volume?

**Solution:** We're told that $\left|\frac{\Delta r}{r}\right| < 0.02$, and $\left|\frac{\Delta h}{h}\right| < 0.005$, and we want to know about $\left|\frac{\Delta V}{V}\right|$. Well,

$$\frac{\Delta V}{V} \approx \frac{dV}{V} = \frac{2\pi r h dr + \pi r^2 dh}{\pi r^2 h} = 2\frac{dr}{r} + \frac{dh}{h}.$$

Therefore, $\left|\frac{\Delta V}{V}\right| \approx \left|2\frac{dr}{r} + \frac{dh}{h}\right| \leq 2\left|\frac{dr}{r}\right| + \left|\frac{dh}{h}\right| < 2(0.02) + 0.005 = 0.045.$

[3] You might wonder if things always work out this neatly; in this problem we didn't even need to specify the values of the various variables in order to answer a question about their relative rates of change. In fact this will be the case whenever the original formula is composed of products and quotients; if there is a sum or a difference involved then we'll need to specify the values of the variables.

**Geometric Interpretation of the Differential**

For a better idea of why the differential has the structure that it does, consider the figure below.

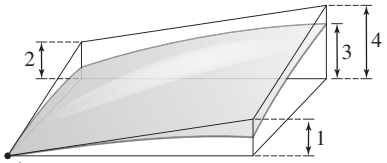



Figure 1:

In this figure, the shaded surface represents part of the graph of a function $f(x, y)$. Below it lies the $xy$-plane, and above it is the tangent plane to $f$, centered at the lower left corner, which we'll take to be the point $(x_0, y_0)$. Also let the dimensions of the rectangle in the $xy$-plane be $\Delta x$ and $\Delta y$. Now, consider the cross-section of the surface where $y = y_0$ (the front face of the diagram). The distance labelled "1" is the distance that the tangent line to $f(x, y_0)$ rises as $x$ increases from $x_0$ to $x_0 + \Delta x$. From your understanding of single-variable calculus, you should recognize that this is given by $\dfrac{\partial f}{\partial x}\Delta x$ ("rise"="slope"×"run"). Next, consider the cross section where $x = x_0$ (the left face). The distance labelled "2" is the distance that the tangent line to $f(x_0, y)$ rises as $y$ increases from $y_0$ to $y_0 + \Delta y$. This is given by $\dfrac{\partial f}{\partial y}\Delta y$. Finally, since the tangent plane is flat, the distance labelled "4" must be the sum of these two distances: the distance that the tangent plane rises as we move from the point $(x_0, y_0)$ to the point $(x_0 + \Delta x,\ y_0 + \Delta y)$ is $df = \dfrac{\partial f}{\partial x}\Delta x + \dfrac{\partial f}{\partial y}\Delta y$. Provided that the distances $\Delta x$ and $\Delta y$ are small enough, this quantity $df$ should serve as a useful approximation to the distance labelled "3". That's $\Delta f$: the change in the value of the the original function $f(x, y)$ as we move between these same two points.

# 4 Introduction to Vector Functions: Parametric Representations of Curves

Early in Math 117 we defined the sine and cosine functions as the coordinates of a point moving around the unit circle; in Figure 2 the point labelled $(x, y)$ has coordinates given by the equations

$$\begin{aligned} x &= \cos(\omega t) \\ y &= \sin(\omega t) \end{aligned} \qquad t \in [0, 2\pi/\omega] \tag{4}$$

where $\omega t$ is the angle, as measured from the positive $x$-axis.

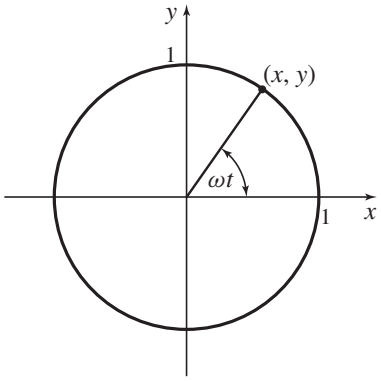



Figure 2:

This is, in fact, a simple example of a very useful way of describing curves in general. Instead of stating a relationship between $x$ and $y$, we may describe each of $x$ and $y$ as a function of a third variable (the "parameter"). The two equations labelled (4) are referred to as *parametric equations* for the unit circle.

If we think of the parameter $t$ as representing time, then we can imagine $x$ and $y$ as representing the coordinates of a point which is moving about in the $xy$-plane, and tracing out a curve as it does so. The advantage of this approach is that regardless of whether $y$ is a function of $x$ or $x$ is a function of $y$, we can always treat $x$ *and* $y$ as functions of time. This can save us from having to break problems down into separate cases (as we would otherwise often have to do for the unit circle, with $y$ being given by either $\sqrt{1-x^2}$ or $-\sqrt{1-x^2}$). Note, however, that we've actually introduced more information; while a relationship between $x$ and $y$ describes a curve, our parametric equations also indicate the direction in which the curve is to be traced out, and at what speed. Also, by placing restrictions on the values of $t$ to be considered, we can specify particular sections of these curves. For this reason, we make a

distinction between the *curve* and the *path*. For example, for the parametric equations above (Equation 4), the path is the entire unit circle, traversed counterclockwise, starting from the point $(1, 0)$, with angular frequency $\omega$. We could describe exactly the same circle by using the parametric equations

$$\begin{aligned} x &= \sin(t) \\ y &= \cos(t) \end{aligned} \qquad t \in [0, 3\pi].$$

The curve is the same, but this time the path starts at $(0, 1)$, moves clockwise, has angular frequency 1 Hz (if $t$ is measured in seconds), and travels around the circle one and a half times.

Working with parametric equations is easiest if we think of them together as single entities, with a scalar input and a vector output. That is, they can be viewed as *vector functions*, which we'll write as $\vec{r}(t)$, with $\vec{r} : \mathbb{R} \to \mathbb{R}^2$.

$$t \to \boxed{\vec{r}} \to \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}.$$

The vector $\vec{r}(t)$ is interpreted as being based at the origin, so that the tip of the vector traces out the path. With this perspective, we can then define the derivative of such a function simply as $\vec{r}\,'(t) = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}$. If we continue to think of $\vec{r}(t)$ as describing the motion of a point, then $\vec{r}\,'(t)$ is its velocity, while the speed is given by $||\vec{r}\,'(t)||$. Naturally, we can also interpret $\vec{r}\,''(t)$ as being the acceleration.[4]

**Example:** Imagine a particle moving in a circle of radius $a$, according to the vector function $\vec{r}(t) = \begin{bmatrix} a\cos(\omega t) \\ a\sin(\omega t) \end{bmatrix}$. Its velocity is given by

$$\vec{v}(t) = \vec{r}\,'(t) = \begin{bmatrix} -a\omega\sin(\omega t) \\ a\omega\cos(\omega t) \end{bmatrix},$$

so its speed is

$$||\vec{v}(t)|| = \sqrt{a^2\omega^2\sin^2\omega t + a^2\omega^2\cos^2\omega t} = a\omega.$$

[4] As you may already be aware, the traditional way to denote vectors is to use boldface type ($\mathbf{r}(t)$), and to use arrows only when writing by hand ($\vec{r}(t)$). I've decided to use arrows throughout; I just think they're easier to see!

That is, we have constant *speed*, even though the direction of motion keeps changing. Also, the acceleration is

$$\vec{a}\,(t) = \vec{r}\,''\,(t) = \begin{bmatrix} -a\omega^2 \cos \omega t \\ -a\omega^2 \sin \omega t \end{bmatrix}.$$

If you look carefully at this, you'll see that $\vec{a}\,(t) = -\omega^2 \vec{r}\,(t)$. That is, the acceleration vector has exactly the opposite direction of the position vector; it points directly back towards the origin. What we've discovered is the mathematical description of centripetal acceleration!

**Parameterization of Other Types of Curves**

Later in your studies you will discover that we occasionally find it necessary to describe given curves in parametric form. For curves in two dimensions this is not difficult; the following examples should be enough to give you the idea.

**Examples:**

a) Write down parametric equations for the ellipse with equation $x^2 + 4y^2 = 4$.

**Solution:** An ellipse is not very different from a circle; the only difference is in the coefficients of $x$ and $y$. Therefore we can parameterize it in a similar fashion... we just need to adjust those coefficients. Since we want the values of $x$ to range between $-2$ and 2, while the values of $y$ should range between $-1$ and 1, we can try letting

$$\vec{r}\,(t) = \begin{bmatrix} x\,(t) \\ y\,(t) \end{bmatrix} = \begin{bmatrix} 2\cos \omega t \\ \sin \omega t \end{bmatrix}, \quad t \in [0, 2\pi/\omega]\,.$$

We can easily verify that this does indeed give us the correct relationship between coordinates:

$$x^2 + 4y^2 = 4\cos^2 \omega t + 4\sin^2 \omega t = 4.$$

Note that we can choose any value we wish for $\omega$ (this determines the speed of motion, but doesn't change the curve).

b) Write down parametric equations for the ellipse with equation $(x-1)^2 + 4\,(y-2)^2 = 4$.

**Solution:** This is the same ellipse as in (a), but with its center shifted to the point $(1, 2)$. Such a shift is easy to incorporate into the parametric equations; we need simply

$$\vec{r}\,(t) = \begin{bmatrix} 2\cos\omega t + 1 \\ \sin\omega t + 2 \end{bmatrix}, \qquad t \in [0, 2\pi/\omega]\,.$$

c) Write down parametric equations for the section of parabola with equation $y = x^2$, with $x \in [-2, 2]$.

**Solution:** This is almost a trick question; in a sense the curve is *already* parameterized, with $x$ as the parameter! If we insist on using $t$ as the parameter, then all we have to do is let $x = t$, and let $y = x^2 = t^2$. That is, as a vector function we have

$$\vec{r}\,(t) = \begin{bmatrix} t \\ t^2 \end{bmatrix}, \quad t \in [-2, 2]\,.$$

Notice that there's really nothing special about the parabola; we can do this with *any* curve described explicitly as $y = f\,(x)$; we just let $\vec{r}\,(t) = [t, f\,(t)]$.

d) Write down parametric equations for the straight line segment from $(2, 3)$ to $(4, 7)$.

**Solution:** We could proceed by finding the equation of this line, and then using the idea of example (c). However, there is a quicker way, and the task of parameterizing straight line segments is one we'll have to perform quite often, so this will be useful. The idea is that, for a line, we should have a constant velocity vector (at least, its direction must be constant), and if we are to move from $(2, 3)$ to $(4, 7)$, we want it to have direction $(2, 4)$. That is, we want $\vec{r}\,'\,(t) = (2, 4)$ (or we could use any multiple of this vector). So, let's set $x'\,(t) = 2$ and $y'\,(t) = 4$. Antidifferentiating gives $x\,(t) = 2t + C_1$ and $y\,(t) = 4t + C_2$. In vector form, we then have

$$\vec{r}\,(t) = (2t + C_1, 4t + C_2)$$

which can be expressed as

$$\vec{r}\,(t) = (2, 4)\,t + (C_1, C_2)\,.$$

What should the two constants of integration be? Well, if we set $t = 0$, we see that $(C_1, C_2)$ can be any point on the line segment; for simplicity we might as well take it to be our initial point $(2, 3)$. You can then see that setting $t = 1$ takes us to the endpoint $(4, 7)$:

$$\vec{r}\,(t) = (2, 3) + (2, 4)\, t, \qquad t \in [0, 1]\,.$$

We could write this as $\vec{r}\,(t) = (2 + 2t, 3 + 4t)$, or

$$\vec{r}\,(t) = \begin{bmatrix} 2 + 2t \\ 3 + 4t \end{bmatrix}, \qquad t \in [0, 1]\,,$$

if you prefer one of those notations. In general, then, for a line segment from $(x_0, y_0)$ to $(x_1, y_1)$ we simply set

$$\vec{r}\,(t) = (x_0, y_0) + (x_1 - x_0, y_1 - y_0)\, t, \qquad t \in [0, 1]\,.$$

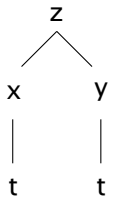
**Comment:** Everything in this discussion can easily be generalized for curves in $\mathbb{R}^n$. For example, a vector function $\vec{r}\,(t) = \begin{bmatrix} x\,(t) \\ y\,(t) \\ z\,(t) \end{bmatrix}$ would describe the path of a particle moving in three dimensions, with velocity $\vec{r}\,'\,(t)$. It is also possible to describe *surfaces* in parametric form (we need *two* parameters), but we'll leave that discussion for a later course.

# 5 The Basic Chain Rule, the Gradient Vector, and Directional Derivatives

## 5.1 The Basic Chain Rule

Consider a function $f(x, y)$, where $x$ and $y$ are themselves functions of a parameter $t$. That is, suppose $f = f(\vec{r})$, where $\vec{r}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$. In this circumstance we can view the output as a function of $t$. We can represent the relationship between the variables with a simple "tree diagram":

Figure 3:




This idea will be useful later for generalizing the rule for cases in which $f$ depends on more than two variables, or where $x$ and $y$ each depend on more than one, etc.

**Example 1:** Suppose $f(x, y) = x^2 e^y$, with $x = t^2 - 1$ and $y = \sin t$. Then we may write *either*

$$z = f(x, y) = x^2 e^y$$

$$\text{or} \quad z = g(t) = \left(t^2 - 1\right) e^{\sin t}.$$

To find $\dfrac{dz}{dt}$, we could just differentiate after substitution (that is, calculate $\dfrac{dz}{dt}$ as $g'(t)$), but there is another option. The differential reads

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy,$$

which we could equally well write as

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy,$$

since $z = f(x, y)$. Dividing by the infinitesimal $dt$ gives

$$\boxed{\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}} \tag{5}$$

This is our most basic form of the Chain Rule for multivariate calculus.[5] It may be referred to as the *Chain Rule for Paths*.

**Example 2:** Let $f$ be the function given in Example 1, and let $z = f(x, y) = g(t)$. Find the value of $\left.\frac{dz}{dt}\right|_{t=0}$.

**Solution:** We have $\frac{\partial z}{\partial x} = 2xe^y$, $\frac{\partial z}{\partial y} = x^2e^y$, $\frac{dx}{dt} = 2t$, and $\frac{dy}{dt} = \cos t$. Therefore

$$\frac{dz}{dt} = (2xe^y)(2t) + \left(x^2e^y\right)(\cos t).$$

Now, to obtain $\frac{dz}{dt}$ as a function of $t$, we'd need to substitute the given functions for $x$ and $y$ into this expression. However, since we're asked for the derivative *at* $t = 0$, this isn't necessary. When $t = 0$, we find that $x = -1$ and $y = 0$, so we can put all three of these values into our derivative: we find that $\left.\frac{dz}{dt}\right|_{t=0} = 1$.

## 5.2 The Gradient Vector

Now, let's return to our discussion of scalar fields $f : \mathbb{R}^2 \to \mathbb{R}$. We've seen that these have two derivatives. However, the two derivatives usually appear together in our calculations, and there's a useful way to join them together. We define the *gradient vector* of $f(x, y)$ as

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right).$$

[5] You might be wondering why we switched to using $dz$ instead of $df$. The reason is that the name $f$ refers to a function of $x$ and $y$ (in our example it refers to the rule $x^2e^y$). Strictly speaking, this cannot be differentiated with respect to $t$. The name $z$, on the other hand, refers to the output quantity, which can be considered to be dependent on either $x$ and $y$ or on $t$. If we wish to write the Chain Rule using the names of the functions instead of the variables (which is sometimes desirable), then we should write it as

$$\frac{dg}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt},$$

where $g$ is the function of $t$ and $f$ is the function of $x$ and $y$.

The symbol $\nabla$ was introduced by Sir William Rowan Hamilton, but unfortunately he neglected to give it a name! As a consequence, several different pronunciations have come into common usage; $\nabla f$ can be read as "grad $f$", or "del $f$", or "nabla $f$". Whatever we choose to call it, it allows us to write our results more concisely, by using vectors throughout:

- The linear approximation $f(x,y) \approx f(a,b) + f_x(a,b) \cdot (x-a) + f_y(a,b) \cdot (y-b)$ can now be written as

$$f(\vec{r}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{r} - \vec{a})$$

  where $\vec{r} = (x,y)$ and $\vec{a} = (a,b)$ (and the symbol $\cdot$ now denotes the dot product).

- The basic chain rule can also be rewritten:

$$\begin{aligned} \frac{dz}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}\right) \\ &= \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t) \end{aligned}$$

  Compare the structure of the above results to the single-variable versions: our linear approximation looks a lot like the familiar formula $f(x) \approx f(a) + f'(a) \cdot (x-a)$, and the chain rule looks a lot like the familiar formula $\dfrac{dy}{dt} = f'(g(t)) \cdot g'(t)$. In essence, the gradient vector $\nabla f$ acts as *the* derivative of $f(x,y)$, so there's a good reason for calling $f_x$ and $f_y$ the "partial" derivatives!

## 5.3 Directional Derivatives

Now that we have the multivariate version of the chain rule, we can start to answer some more sophisticated questions. For example, we introduced $f_x$ and $f_y$ as the rates of change of $f$ as we move in the direction of increasing $x$ or $y$, respectively, but what if we're interested in the rate of change of $f$ as we move in some other direction? For this we have a precise definition:

**Definition:** The *directional derivative* of $f(x,y)$ in the direction of a unit vector $\vec{u} = (u_1, u_2)$ at the point $\vec{a} = (a,b)$ is denoted by $D_{\vec{u}} f(a,b)$, and defined as

$$D_{\vec{u}} f(a,b) = \lim_{h \to 0} \frac{f(\vec{a} + h\vec{u}) - f(\vec{a})}{h} = \lim_{h \to 0} \frac{f(a + hu_1, b + hu_2) - f(a,b)}{h}.$$

We hope that you can see why this is the definition we want[6], but clearly we'd like to have a more practical formula. In fact, there is one, and its derivation is based on a clever observation. The expression $f(a+hu_1, b+hu_2)$ is a function of $h$, since $a$, $b$, $u_1$, and $u_2$ are all constants. Let's call this function $g$:

$$g(h) = f(a+hu_1, b+hu_2).$$

Furthermore, *our directional derivative $D_{\vec{u}}f(a,b)$ is the derivative of this function at $h=0$!* To see this, just consider the definition of the ordinary derivative:

$$g'(0) = \lim_{h\to 0} \frac{g(h)-g(0)}{h} = \lim_{h\to 0} \frac{f(a+hu_1, b+hu_2) - f(a,b)}{h}$$

$$= D_{\vec{u}}f(a,b).$$

Now, we've just seen that in situations like this, we have an alternative way of calculating the derivative; we can use the chain rule! Using $g(h) = f(x(h), y(h))$, where $x(h) = a+hu_1$ and $y(h) = b+hu_2$, we find

$$g'(h) = \frac{\partial f}{\partial x}\frac{dx}{dh} + \frac{\partial f}{\partial y}\frac{dy}{dh}$$

$$= \frac{\partial f}{\partial x}u_1 + \frac{\partial f}{\partial y}u_2$$

$$= \nabla f(x,y)\cdot\vec{u},$$

and setting $h=0$ tells us that $g'(0) = \nabla f(a,b)\cdot\vec{u}$. Therefore

$$\boxed{D_{\vec{u}}f(a,b) = \nabla f(a,b)\cdot\vec{u}.}$$

**Example:** For $f(x,y) = x^2+y^2$, find the slope at the point $(1,-1)$ in the direction of the vector $(3,4)$.[7]

[6] The numerator of the quotient represents the change in the value of $f$ as we move a distance $h$ from the point $\vec{a}$, in the direction $\vec{u}$. The quotient represents the average change of $f$ along this path, and the entire expression is the limit of this as the distance approaches zero. That is, it's exactly the same concept as the derivative we're familiar with, but set in a specific direction in two dimensions.

[7] We could phrase this differently; we could say "in the direction of the *point* $(4,3)$". We would then have to calculate the direction vector as $(4,3)-(-1,1) = (3,4)$.

**Solution:** First note that $(3,4)$ is not a unit vector, so before we can use our formula we must normalize it:

$$\vec{u} = \frac{(3,4)}{\sqrt{3^2+4^2}} = \left(\frac{3}{5}, \frac{4}{5}\right).$$

Now, $\nabla f\,(x,y) = (f_x, f_y) = (2x, 2y)$, so $\nabla f\,(1,-1) = (2,-2)$. Therefore

$$D_{\vec{u}} f\,(1,-1) = \nabla f \cdot \vec{u} = (2,-2) \cdot \left(\frac{3}{5}, \frac{4}{5}\right) = -\frac{2}{5}.$$

### Interpretation of the Gradient Vector Itself

We have now seen a couple of ways in which the gradient vector can be used, but you might be wondering what it *is*. Recall that $\vec{a} \cdot \vec{b} = |\vec{a}|\ |\vec{b}| \cos\theta$, where $\theta$ is the angle between the vectors $\vec{a}$ and $\vec{b}$. If we apply this to our formula for the gradient vector, we find that $\nabla f \cdot \vec{u} = ||\nabla f||\ ||\vec{u}|| \cos\theta$. Since $\vec{u}$ is a *unit* vector, though, this is just

$$\nabla f \cdot \vec{u} = ||\nabla f|| \cos\theta.$$

This means that $D_{\vec{u}} f$ has its maximum value when $\theta = 0$; that is, when $\vec{u}$ is in the same direction as $\nabla f$. Furthermore, that maximum value is $||\nabla f||$. In other words, at any given point, the direction and magnitude of the steepest slope of the graph of $f$ are given by the vector $\nabla f$.[8]

**Note:** Since $\nabla f$ represents a different vector at each point, it is in fact our first example of a *vector field*; $\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$ (a single "function" which takes multiple inputs and gives multiple outputs; a mapping assigning vectors to vectors). This has a useful graphical interpretation. Consider a contour plot of a function $f$. Since the slope along the level curves is zero (by definition - they are *level* curves!), it can be shown that the gradient vector at any point will always be orthogonal to the level curve passing through that point. Therefore, if we know what the contour plot looks like, then we know what the gradient field looks like (at least in principle), and vice versa. This means that we have a choice; if we want a two-dimensional representation of the graph of a function $f\,(x,y)$, we could use *either* the contour plot or the

[8] We might also say simply that $\nabla f$ points in the direction of greatest increase of $f$, but keep in mind that it has only two components (that is, it lies in the $xy$-plane). If you imagine yourself standing on the surface that is the graph of $f$, you could think of $\nabla f$ as giving you a "compass direction", but instead of pointing north it tells you which way to head if you want to go directly uphill.

gradient field.

**Example:**

a) Consider the function $f(x, y) = x^2 + y^2$. We've already discussed the contour plot for this function (the contours are concentric circles, and the graph of $f$ is a paraboloid). Now, the gradient of $f$ is $\nabla f = (2x, 2y)$; every one of these vectors is directed away from the origin, with magnitude proportional to the distance from the origin. If we plot a selection of these vectors, and remember that each vector represents the slope of the surface at that point, we get our alternative representation of the paraboloid (see Figure 4).

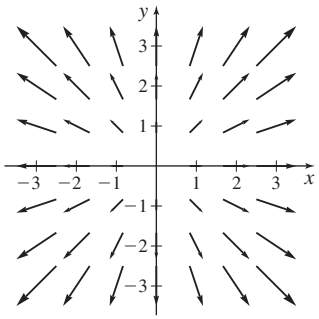



Figure 4:

b) Consider the function $f(x, y) = x^2 - y^2$. You might also recall this one; its graph is the prototypical saddle surface. This time $\nabla f = (2x, -2y)$, and some of these vectors are shown in Figure 5.

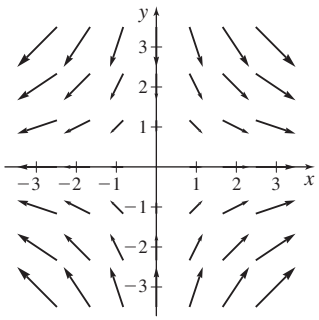


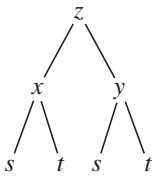

Figure 5:

## 6 Other Forms of the Chain Rule

In our introduction to the multivariate chain rule, we considered the situation in which $z = f(x, y)$, with $x$ and $y$ each being dependent upon a third variable $t$. However, we may encounter problems in which $x$ and $y$ are themselves given by multivariate functions.

**Chain Rule for Mappings:** Perhaps the second most-commonly encountered situation is this: we may have $z = f(x, y)$, where $x = g_1(s, t)$ and $y = g_2(s, t)$. These equations can be interpreted as defining a mapping (or *transformation*) from one coordinate system to another.[9] We can illustrate the relationships between the variables this way:

Figure 6:




In this situation we may view $z$ as being dependent upon $s$ and $t$, with partial derivatives given by

$$\begin{aligned} \frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\ \text{and} \quad \frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}. \end{aligned} \tag{6}$$

(Why? To calculate $\dfrac{\partial z}{\partial s}$, all we have to do is view $t$ as constant, and use our previous version of the chain rule! The only difference is that whereas the required derivatives of $x$ and $y$ used to be *ordinary* derivatives, they are now *partial* derivatives.)

This particular version of the chain rule will be an essential tool in later courses.

**Example 1:** Suppose $z = f(x, y)$. If we introduce the change of variables

$$\begin{aligned} x &= \rho\cos\phi \\ y &= \rho\sin\phi, \end{aligned}$$

then we can express $z_{\rho\rho}$ in terms of the derivatives $z_x$ , $z_y$, $z_{xx}$, etc. as follows:

[9] Consider, for example, the equations for converting from Cartesian to polar coordinates, which we discussed briefly in Math 117: we let $x = \rho\cos\phi$ and $y = \rho\sin\phi$.

- To get $z_{\rho\rho}$ we must start with $z_\rho$ (which just requires using Equation (12) above :

$$z_\rho = \frac{\partial z}{\partial \rho} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial \rho} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \rho} = z_x \cos\phi + z_y \sin\phi. \tag{7}$$

- Now we just have to differentiate this result with respect to $\rho$ :

$$\begin{aligned} z_{\rho\rho} = \frac{\partial^2 z}{\partial \rho^2} &= \frac{\partial}{\partial \rho}\left(\frac{\partial z}{\partial \rho}\right) = \frac{\partial}{\partial \rho}\left(z_x \cos\phi + z_y \sin\phi\right) \\ &= \cos\phi \frac{\partial}{\partial \rho}(z_x) + \sin\phi \frac{\partial}{\partial \rho}(z_y) \\ &= \cos\phi \left[\frac{\partial z_x}{\partial x}\frac{\partial x}{\partial \rho} + \frac{\partial z_x}{\partial y}\frac{\partial y}{\partial \rho}\right] + \sin\phi \left[\frac{\partial z_y}{\partial x}\frac{\partial x}{\partial \rho} + \frac{\partial z_y}{\partial y}\frac{\partial y}{\partial \rho}\right] \\ &= \cos\phi \left[z_{xx}\cos\phi + z_{xy}\sin\phi\right] + \sin\phi \left[z_{yx}\cos\phi + z_{yy}\sin\phi\right] \\ &= z_{xx}\cos^2\phi + 2z_{xy}\sin\phi\cos\phi + z_{yy}\sin^2\phi \end{aligned} \tag{8}$$

**Comments:**

1. These expressions appear to have a mixture of functions of $(x, y)$ and $(\rho, \phi)$, which we should avoid, but we also replace the variables within the functions $z_x$ and $z_y$. That is, $z_\rho = z_x(x, y)\cos\phi + z_y(x, y)\sin\phi = z_x(\rho\cos\phi, \rho\sin\phi)\cos\phi + z_y(\rho\cos\phi, \rho\sin\phi)\sin\phi$, so $z_\rho$ is indeed a function of $\rho$ and $\phi$.

2. In the first three lines of the second calculation (8), I've used both the Leibniz and subscript notations together. You should be aware that this is not usually done. Most authors will pick one notation and stick with it, and the most common choice for these calculations is the Leibniz notation. Using this exclusively, (8) looks like this:

$$\begin{aligned} \frac{\partial^2 z}{\partial \rho^2} &= \frac{\partial}{\partial \rho}\left(\frac{\partial z}{\partial \rho}\right) = \frac{\partial}{\partial \rho}\left(\frac{\partial z}{\partial x}\cos\phi + \frac{\partial z}{\partial y}\sin\phi\right) \\ &= \cos\phi \frac{\partial}{\partial \rho}\left(\frac{\partial z}{\partial x}\right) + \sin\phi \frac{\partial}{\partial \rho}\left(\frac{\partial z}{\partial y}\right) \\ &= \cos\phi \left[\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right)\frac{\partial x}{\partial \rho} + \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)\frac{\partial y}{\partial \rho}\right] \end{aligned}$$

$$+\sin\phi\left[\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)\frac{\partial x}{\partial \rho}+\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right)\frac{\partial y}{\partial \rho}\right] \tag{9}$$

$$=\cos\phi\left[\frac{\partial^2 z}{\partial x^2}\cos\phi+\frac{\partial^2 z}{\partial y\partial x}\sin\phi\right]$$

$$+\sin\phi\left[\frac{\partial^2 z}{\partial x\partial y}\cos\phi+\frac{\partial^2 z}{\partial y^2}\sin\phi\right]$$

$$=\cos^2\phi\frac{\partial^2 z}{\partial x^2}+2\cos\phi\sin\phi\frac{\partial^2 z}{\partial x\partial y}+\sin^2\phi\frac{\partial^2 z}{\partial y^2}.$$

In my experience most students find this calculation overwhelming when they first see it in this notation, and my hope is that the notation in (8) makes it clearer that *all we are doing is using the basic rule (6) twice*! The functions $z_x$ and $z_y$ are functions of $x$ and $y$, just as the original function $z$ is, and so if we need to differentiate them with respect to $\rho$ the procedure is exactly the same. However, you should convince yourself that the meaning of each line in (9) is exactly the same as the corresponding line in (8).

### Non-Standard Forms of the Chain Rule

If we encounter a relationship between variables which doesn't fit the two forms of the Chain Rule we've discussed so far, we'll need to determine what form the Chain Rule should take. For this, we can use the "tree diagrams" as a guide: identify all of the routes leading to the desired variable, multiply the derivatives along that route, and add the results.

**Example 2:** Suppose that $w = f(x, y, z)$, with $x = x(s, t)$, $y = y(s)$, and $z = z(t)$ (using $x$, $y$, and $z$ for both variables and functions, to avoid having to use too many letters of the alphabet at once). The appropriate tree diagram is given in Figure 7:

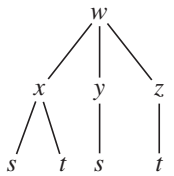



Figure 7:

In this situation we may view $w$ as dependent upon $s$ and $t$, and the partial derivatives are given by

$$w_s = \frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{dy}{ds} = w_x x_s + w_y y'(s)$$

$$w_t = \frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial z}\frac{dz}{dt} = w_x x_t + w_z z'(t)$$

In some of these examples, we discover one good reason for having different notations for ordinary and partial derivatives:

**Example 3:** Suppose that $z = f(x, y)$, where $y = g(x)$. In this case we can either view $z$ as being dependent upon both $x$ and $y$ , or we can view it as being dependent upon $x$ alone (i.e. $z = h(x) = f(x, g(x))$. The tree diagram for this relationship looks like this:

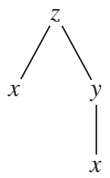



Figure 8:

The derivative of the function of $x$ alone is

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}. \tag{10}$$

You should be able to see that $\frac{dz}{dx}$ and $\frac{\partial z}{\partial x}$ represent *different functions* here; the first is the derivative of the single-variable function we've called $h(x)$, while the second is one of the partial derivatives of the function $f(x, y)$.

Perhaps this will be clearer if we make the example more specific. Suppose $z = x^2 + y^2$, where $y = \sin x$ (so our function $f(x, y)$ is $x^2 + y^2$, while $g(x)$ is $x^2 + \sin^2 x$). Then $\frac{dz}{dx}$ can be calculated in two ways:

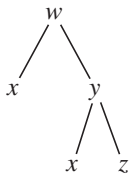
1. Differentiate after substitution: $z = g(x) = x^2 + \sin^2 x$, so $\dfrac{dz}{dx} = g'(x) = 2x + 2\sin x \cos x$ (using the ordinary single-variable chain rule).

2. Use the appropriate multivariate chain rule (our Equation (10)): $\dfrac{\partial z}{\partial x}$ is $\dfrac{\partial f}{\partial x} = 2x$ (and $\dfrac{\partial z}{\partial y} = \dfrac{\partial f}{\partial y} = 2y$), so $\dfrac{dz}{dx} = 2x + 2y\cos x$. But $y = \sin x$, so this is indeed $2x + 2\sin x \cos x$, as expected!

**Note:** We've chosen specific functions here to illustrate the distinction between $\frac{dz}{dx}$ and $\frac{\partial z}{\partial x}$, but if we *know* all of the functions involved then these generalized versions of the chain rule are unnecessary; we can do everything we need to with the single-variable version. Keep in mind that the reason we have these rules is for examples like Example 1, in which we do not (yet) know the functions!

**Example 4:** Suppose $w = x^3 + 2y$, where $y = x \sin z$. In this case we may view $w$ either as being dependent upon $x$ and $y$, or as being dependent upon $x$ and $z$:

Figure 9:




Using the chain rule to find the derivatives of $w$ with respect to $x$ and $z$, we have

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial x} \tag{11}$$

and

$$\frac{\partial w}{\partial z} = \frac{\partial w}{\partial y}\frac{\partial y}{\partial z} \tag{12}$$

but we've run into a problem with our notation; we've got two *different* quantities labelled as $\frac{\partial w}{\partial x}$ in equation (11)! They are both partial derivatives, but the first one is a derivative of $w$ seen as a function of $x$ and $z$, while the second one is a derivative of $w$ seen as a function of $x$ and $y$. Unfortunately our usual notation can't distinguish between these, so to fix this we must either include the variables explicitly, and write

$$\frac{\partial w}{\partial x}(x, z) = \frac{\partial w}{\partial x}(x, y) + \frac{\partial w}{\partial y}(x, y)\frac{\partial y}{\partial x}(x, z),$$

or else we could take care to use names for the functions instead of the variables. Let's try using the name $f$ for the function which gives $w$ from $x$ and $y$: $\boxed{w = f(x, y)}$
and using the name $h$ for the function which gives $w$ from $x$ and $z$: $\boxed{w = h(x, z)}$.

Then we have

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial x}$$

and

$$\frac{\partial h}{\partial z} = \frac{\partial f}{\partial y}\frac{\partial y}{\partial x}.$$

**Example 5:** Verify that for *any* twice-differentiable functions $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$, the function $f(x,t) = g(x+ct) + h(x-ct)$ is a solution to the *wave equation*:

$$\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2}.$$

**Solution:** This is, in the current context, a bit of a trick question. We've been discussing generalizations of the chain rule, and in this case we could come up with a quite complicated-looking one[10]. However, the unknown functions here are (as we've carefully stated), *single-variable functions*! Therefore we don't really need a multivariate chain rule at all. To calculate the partial derivatives of $f$, just treat one variable as constant, and differentiate:

$$f_x = g'(x+ct) + h'(x-ct)$$

$$f_{xx} = g''(x+ct) + h''(x-ct)$$

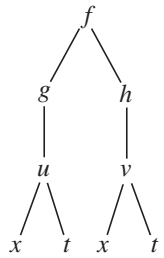
$$f_t = cg'(x+ct) - ch'(x-ct)$$

$$f_{tt} = c^2 g''(x+ct) + c^2 h''(x-ct).$$

Comparing, we find that indeed, $f_{tt} = c^2 f_{xx}$.

[10]What should this rule look like? For simplicity, let's allow ourselves to use $f$, $g$, and $h$ not just as the names for functions, but also for the variables they give as output (as we did with $x$, $y$, and $z$ in Example 2). We'll also need to give names to the input variables for $g$ and $h$: let $u = x + ct$, and let $v = x - ct$. Then we have $f = g + h$, where $g = g(u)$ and $h = h(v)$, while $u = x + ct$ and $v = x - ct$. The tree diagram is given in Figure (10).The appropriate chain rule for $f_x$, then, is $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial g}\frac{dg}{du}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial h}\frac{dh}{dv}\frac{\partial v}{\partial x}$. You should be able to see at a glance that this does indeed reduce to $g'(u) + h'(v)$, since we have explicit formulas for $f$, $u$, and $v$.

Figure 10:



# 7 Optimization Techniques

## 7.1 Unconstrained Optimization

We now turn to the problem of identifying maxima and minima of functions of two variables. First, we need to define them:

A function $f(x, y)$ has a *local maximum* at $(x_0, y_0)$ if $f(x_0, y_0) \geq f(x, y)$ for all $(x, y)$ in some disk centered at $(x_0, y_0)$.[a]

[a] The disk may be small; the idea is that it should be possible to identify a disk small enough so that the inequality holds throughout.

We can state a similar definition for a *local minimum*; for this we need $f(x_0, y_0) \leq f(x, y)$.

If we ask ourselves how we should go about finding these, the first step should be clear; such extrema should only occur at points where the tangent plane is horizontal (or at points where no tangent plane can be defined). If the tangent plane is horizontal, then $\nabla f = \vec{0}$. Therefore we can define a *critical point* as a point at which either *both* $f_x$ and $f_y$ are zero, or else one of them is undefined.

As in the single-variable case, however, a critical point need not correspond to an extremum. Consider, for example, the critical point at the origin on the surface $z = x^2 - y^2$ (the saddle surface). You can check easily that $\nabla f(0, 0) = (0, 0)$, but we know that this point is not an extremum. In fact we borrow our terminology from this example; if a critical point is neither a maximum nor a minimum, we'll call it a *saddle point* (whether the surface is a true saddle shape or not). So, how might we go about determining whether a critical point corresponds to a maximum, a minimum, or a saddle point? In the single-variable case we can look at the sign of the second derivative, to see whether the graph of the given function is concave up or concave down. In multivariate calculus, though, we have *three* second-order derivatives to consider, and the notion of concavity is more complicated. If we consider cross sections of a surface through a saddle point, we will find that some of them are concave up, while others are concave down. To show that we have instead a local maximum [or minimum] we must show the graph to be concave down [or up] along *all* cross-sections through the critical point. This sounds like an imposing task, but there is in fact a simple test we can apply. The proof requires concepts we'll cover in the second half of this course, so we'll simply state the theorem now, and come back to the proof later.

**The Second-Derivative Test for Local Extrema**

Suppose $P_0$ is a critical point of a function $f(x, y)$, and suppose that the second-order partial derivatives of $f$ are continuous in some neighbourhood of $P_0$.
Let $D(x, y) = f_{xx}f_{yy} - (f_{xy})^2$.

- If $D(P_0) > 0$, then $f$ has an extremum at $P_0$.
  - If $f_{xx}(P_0) < 0$ then this extremum is a maximum, whereas if $f_{xx}(P_0) > 0$ then it is a minimum.[a]
- If $D(P_0) < 0$, then $f$ does *not* have an extremum at $P_0$ (that is, it has a saddle point instead).
- If $D(P_0) = 0$, the test gives no conclusion.

[a] We could use either $f_{xx}$ or $f_{yy}$ for this part of the test; at an extremum the concavity will be the same along any cross section.

That third possibility looks distressing; what on earth are we supposed to do if the test fails? Well, we'll have to find some other way to analyze the problem. That *could* be quite difficult, but there are cases in which it is actually really easy. For example, consider the function $f(x, y) = x^4 + y^4$. There's a critical point at $(0, 0)$, and the second-derivative test fails there, but we can tell by inspection that the critical point is a minimum (since $f > 0$ at every point $(x, y) \neq (0, 0)$).

**Example:** Find and classify the local extrema of the function $f(x, y) = x^4 + y^4 - 4xy$.

**Solution:** To locate the critical points, we need to set $\nabla f = \vec{0}$:

$$f_x = 4x^3 - 4y = 4(x^3 - y) = 0$$

$$f_y = 4y^3 - 4x = 4(y^3 - x) = 0$$

The first of these gives us $y = x^3$, and so the second becomes $x^9 - x = 0$, i.e. $x(x^8 - 1) = 0$, so $x = 0$, 1, or $-1$. Since $y = x^3$, we have three critical points: $(0, 0)$, $(1, 1)$, and $(-1, -1)$.

To classify these, we need the second derivatives:

$$f_{xx} = 12x^2, \qquad f_{xy} = 4, \qquad f_{yy} = 12y^2$$

from which we construct the function $D(x, y) = f_{xx}f_{yy} - (f_{xy})^2 = 144x^2y^2 - 16$. We now evaluate this at each of the critical points:

- $D(0, 0) = -16 < 0$, so $(0, 0)$ is a saddle point.
- $D(1, 1) = 128 > 0$, so $(1, 1)$ is a local extremum. Specifically, it is a local minimum, since $f_{xx}(1, 1) = 12 > 0$.
- $D(-1, -1) = 128$, and $f_{xx}(-1, -1) = 12$, also, so $(-1, -1)$ is also a local minimum.

**Comment:** Just as in single-variable calculus, the identification of local extrema can help us with graphing, or at least with producing a contour plot. In the neighbourhood of a local extremum (maximum *or* minimum), the level curves will be (approximately) concentric circles or ellipses, while in the neighbourhood of a saddle point, one level curve will cross over itself (the prototypical examples are the paraboloid $z = x^2 + y^2$ and $z = x^2 - y^2$, which we discussed in our introduction to multivariate calculus). If we consider the example above, this tells us that the parts of the contour plot close to our three critical points are as we see below:

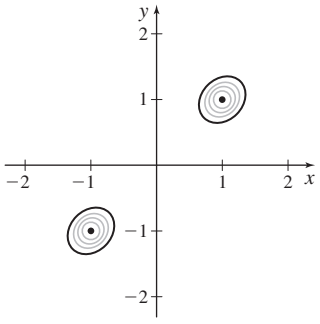



Figure 11:

To fill out the rest of the plot, consider that if we zoom out far enough, the two minima will appear to merge into one (in other words, for large values of $x$ and $y$, we have $f \approx x^4 + y^4$, which has a minimum at the origin). Therefore the level curves far from the origin should be approximately elliptical as well (and if we go far enough out then the level curves will have approximately the form of the curves $x^4 + y^4 = K$, which might be described as rounded

squares). Also, since different level curves cannot intersect, the curve through the origin must look something like a "figure eight"). We then fill in the gaps as sensibly as we can.

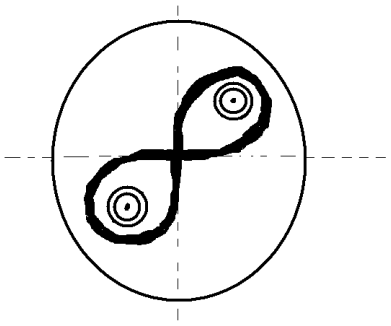


Figure 12:

So far this is just a (very) rough approach by hand, but we're not divorced too far from reality; the result using Maple is shown in Figure 13.

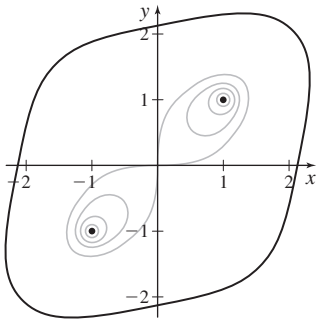



Figure 13:

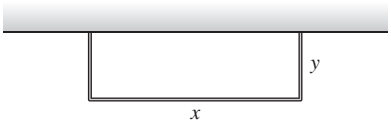
## 7.2 Constrained Optimization

With functions of two variables, there is a second kind of optimization problem that we may encounter. The concept may actually be a familiar one, because there is a famous example (with several variations) which you may have seen in high school:

**Example: The Fencing Problem** Suppose we wish to construct a rectangular enclosure against an existing wall, and we have enough fencing material to span 400m. What's the maximum area than we can enclose, and how should we choose the dimensions of the enclosure in order to achieve it?

**Solution:** If we let $x$ and $y$ be the dimensions of the enclosure, as in Figure 14,

Figure 14:

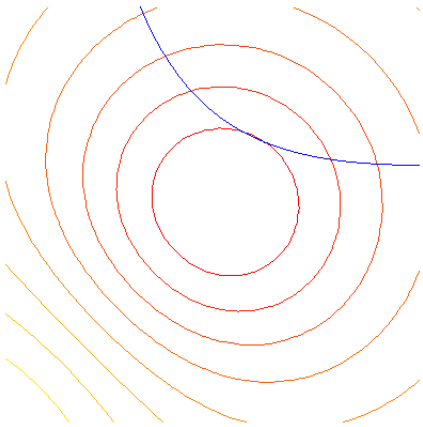



then the problem is to maximize the function $A = xy$, *subject to the constraint* $x+2y = 400$. For simple examples such as this the solution is easy; we can solve the constraint for one of the variables ($x = 400 - 2y$), substitute the result into the problem to be maximized ($A = (400 - 2y)y = 400y - 2y^2$), and then use our single-variable techniques (we find that $A'(y) = 400 - 4y = 4(100 - y)$, so the only critical point is at $y = 100$). We discover that the optimal choices are $x = 200$ and $y = 100$, which gives $A = 20000$ m$^2$.

It will not, however, always be so easy to solve these problems. In particular, it might not be possible to solve the constraint for either one of the variables - and this would prevent us from using the above method. For this reason we now develop a more general method, based upon several of the ideas we have discussed in this course so far.

Suppose we wish to find the local extrema of a function $f(x, y)$, subject to a constraint $g(x, y) = K$ (this is our most general way of expressing a relationship between the two variables $x$ and $y$). Consider the contour plot of $f$, and consider the graph of the constraint curve. These might look something like you see here (the level curves of the function $f$ to be optimized are the roughly circular rings, and the constraint curve is the other one, intersecting some of the rings):

Figure 15:



If we imagine moving along the constraint curve and observing the values of $f$ that we encounter along the way, we realize that ***extreme values of*** $f(x, y)$ ***along the constraint***

***curve will occur when the constraint curve touches a level curve of $f$ tangentially.***

How might we go about locating these points mathematically? You might be able to think of a couple of different strategies, but the simplest result comes from the following observations:

- We know that $\nabla f$ is always orthogonal to the level curves of $f$.
- If we view the constraint curve as being one particular level curve of a function $g(x, y)$, then $\nabla g$ will be orthogonal to the constraint curve.[11]
- Therefore at any point where the constraint curve touches a level curve of $f$ tangentially, it must be that $\nabla g$ is parallel to $\nabla f$ .

That gives us the essence of ***the Method of Lagrange:*** to find the critical points of $f(x, y)$ subject to a constraint $g(x, y) = K$ (where $K$ is a constant), we'll look for the values of $x$ and $y$ for which

$$\nabla f = \lambda \nabla g \qquad \text{and} \qquad g(x, y) = K, \tag{13}$$

for some value of $\lambda$.

We're using the equation $\nabla f = \lambda \nabla g$ to reflect the requirement that the two vectors be parallel. Do we want to allow $\lambda$ to be zero? Yes we do! That will capture the special cases in which $\nabla f = \vec{0}$ at the point in question, meaning that the constraint curve passes directly over a critical point (of the unconstrained variety) of $f(x, y)$. Any local extrema of the unconstrained function will certainly be local extrema for the constrained problem as well.

On the other hand, to reflect the condition that $\nabla g$ and $\nabla f$ be parallel, we could have written $\nabla g = \lambda_1 \nabla f$. Could $\lambda_1$ be zero? (In other words, do we need to consider points at which $\nabla g = \vec{0}$?) This may not be so obvious, but we do indeed have to include this possibility explicitly in our method.[12]

So, here's the full ***the Method of Lagrange***:

[11] For example, in the area problem, we have $g(x, y) = x + 2y$. The level curves of this function have the form $x + 2y = K$, and our constraint $x + 2y = 400$ is simply one of these level curves. The vector $\nabla g = (1, 2)$ is always orthogonal to this line.

[12] If $\nabla g(a, b) = \vec{0}$, then our auxiliary function $g(x, y)$ has a critical point at $(a, b)$. Our constraint curve $g(x, y) = K$ is a level curve of $g$, so what could it look like if it includes a critical point? Well, one possibility is that the critical point of $g$ is a (true) saddle point. That would mean that our constraint curve crosses over itself, and it is indeed possible that the maximum value of $f$ on the curve occurs at that intersection.

To find the critical points of $f(x,y)$ subject to a constraint $g(x,y) = K$ (where $K$ is a constant), find the values of $x$ and $y$ for which

$$\nabla f = \lambda \nabla g \qquad \text{and} \qquad g(x,y) = K \qquad \text{for some constant } \lambda, \tag{14}$$

$$\text{or} \qquad \nabla g = \vec{0} \qquad \text{and} \qquad g\,(x,y) = K.$$

**Example 1:** Revisiting the area problem with our new method, we have $A(x,y) = xy$ and $g(x,y) = x + 2y = 400$, so we have $\nabla A = (y,x)$ and $\nabla g = (1,2)$. Notice that $\nabla g$ is never $(0,0)$. Therefore we just need to solve the system of equations $\nabla A = \lambda \nabla g$, $x + 2y = 400$. This is in fact three equations in three unknowns:

$$y = \lambda$$

$$x = 2\lambda$$

$$x + 2y = 400.$$

In general these systems of equations may be nonlinear, but this one happens to be linear, and an easy example, too: the first two equations give $x = 2y$, so the third one becomes $4y = 400$, and so we find $y = 100$, from which $x = 200$ and $A = 20000$, as expected.

**Notes:**

1. If the constraint is not given in the form $g(x,y) = K$, then we must rewrite it!

2. The proportionality constant $\lambda$ is called a *Lagrange multiplier*. We normally don't care what the value of this is, unless we need to calculate it in order to solve for $x$ and $y$. However, it does have an interesting interpretation; it can be shown that $\lambda$ represents the rate of change of $f$ with respect to changes in the constraint value. For example, in the fencing problem we have $\lambda = 100$, and this means that if we had 401m of fencing instead of 400m, we'd be able to enclose approximately $100\text{m}^2$ more!

**Identifying Absolute Extrema Under Constraints**

It is important to realize that the method of Lagrange merely *locates* critical points; by itself *it does not us whether they are local maxima, minima, or inflection points.* Using calculus to classify them would require a different method; we'd need to parameterize the constraint curve and use the chain rule to find a second derivative of $f$ along the curve. Fortunately, though, we'll usually be more interested in whether the points can be identified as *absolute* maxima or minima, and this is normally a bit easier. The specific steps depend on the nature of the constraint curve:

- If the curve has endpoints, then we must calculate the values of $f(x, y)$ at those points as well, and include them in the comparison.
- If the curve is not closed, and has no endpoints (that is, if it is of infinite extent), then $f(x, y)$ may not even be bounded. We'll need to consider the limits of $f(x, y)$ in the two directions along the curve to determine whether we have any absolute extrema at all.
- Of course, in simple problems this work may be unnecessary. For example, for the fencing problem it should be clear that our calculations give the maximum possible area.

**Extension to Higher Dimensions**

The method works for functions of more than two variables as well, and for such functions it can even be extended for problems with multiple constraints. For example, if we needed to find the critical points of a function $f(x, y, z)$, subject to the constraints $g_1(x, y, z) = K_1$ and $g_2(x, y, z) = K_2$, we'd set $\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$, with $g_1 = K_1$ and $g_2 = K_2$. Such a problem can be interpreted geometrically as finding the critical values of $f(x, y, z)$ along the curve of intersection of the two surfaces defined by the two constraints.

**Example 2:** Determine the dimensions of a rectangular box, open at the top, having a volume of $4\text{m}^2$, requiring the least amount of material for construction (see Figure 16)

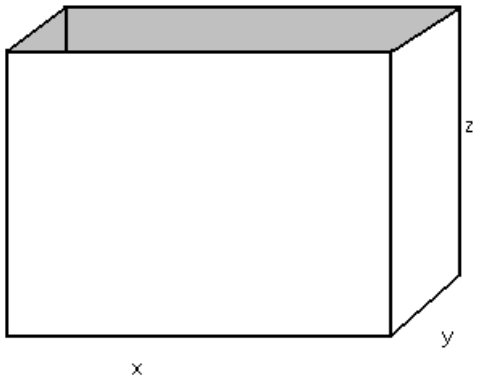



Figure 16:

**Solution:** We wish to minimize the surface area $S(x, y, z) = xy + 2xz + 2yz$ subject to the constraint $V(x, y, z) = xyz = 4$. Setting $\nabla S = \lambda \nabla V$ means setting $(S_x, S_y, S_z) = \lambda(V_x, V_y, V_z)$. Calculating these partial derivatives and coupling these three equations with the constraint gives us four equations in four unknowns:

1. $y + 2z = \lambda yz$
2. $x + 2z = \lambda xz$
3. $2x + 2y = \lambda xy$
4. $xyz = 4$

This system is nonlinear, and there are no standard techniques for solving nonlinear systems of equations. However, for these systems arising from the method of Lagrange we can always start by solving something for $\lambda$; in our current example let's isolate it in all of the first three equations:

$$\begin{aligned} \lambda &= \frac{1}{z} + \frac{2}{y} \\ &= \frac{1}{z} + \frac{2}{x} \\ &= \frac{2}{y} + \frac{2}{x} \end{aligned}$$

We've got three expressions for $\lambda$ here; comparing the first two gives $x = y$ and comparing the second two gives $y = 2z$ $(= x)$. With this information, our fourth equation (the constraint) becomes $(x)(x)(\frac{x}{2}) = 4$, so $x^3 = 8$. Therefore $x = 2$, from which we find $y = 2$ and $z = 1$.

Can $\nabla V$ be $(0,0,0)$? No. The equation $(yz, xz, xy) = (0,0,0)$ is only satisfied at the origin, which does not satisfy the constraint $xyz = 4$.

**A Final Comment:**

The Method of Lagrange is presented differently in many textbooks. For one thing, the possibility that critical points might occur where $\nabla g = \vec{0}$ is usually overlooked. That's simple enough - it's a mistake, but examples in which that possibility turns out to be important are rare.

You may also see the method described in an entirely different way:

*To locate the critical points of $f$ subject to a constraint $g = 0$,*

*define the auxiliary function $\phi(x, y, \lambda) = f(x, y) - \lambda g(x, y)$,*

*and find the critical points of $\phi$.*

This is nothing more than a trick to obtain the same set of equations as above (Equation 14), without having to refer to gradient vectors (this way it's possible to include the method in textbooks and courses intended for students in programs which are less mathematically demanding). There is no other advantage to it, and in fact it leads to one major misunderstanding: it's natural to think that you should be able to classify the critical points by using the second-derivative test on $\phi$, *but this is not the case*! While the critical points of $\phi$ correspond to the critical points we're looking for, the maxima / minima / saddle points *do not*, in general.

# 8 Integration of Scalar Fields

We now turn our attention to the question of how integral calculus might be generalized for functions of more than one variable. It might have occurred to you that there should such a thing as "partial integration", as a counterpart to partial differentiation. This is indeed the case, and in fact we don't even need any new notation; we can simply write $\int f(x,y)dx$ and $\int f(x,y)dy$. The differential identifies the variable of integration for us, so for the former we treat $y$ as a constant, and for the latter we treat $x$ as a constant. The only twist is that what we're used to calling the "constant of integration" need only be constant *with respect to the variable of integration*; it may depend upon the other variable(s), and unless we know otherwise we must write it as such.

**Example:** $\int (x^2+y^2)\,dy = x^2y + \dfrac{y^3}{3} + g(x)$.

We can use this idea to identify a function (up to a constant) when we know both of its partial derivatives.

**Example:** Suppose we know that $f_x = 6xy^2 + e^y$, and $f_y = 6x^2y + xe^y + \sin y$. Then we may write

$$f(x,y) = \int f_x dx = \int \left(6xy^2 + e^y\right) dx = 3x^2y^2 + xe^y + g(y).$$

At the same time, we also know that

$$f(x,y) = \int f_y dy = \int \left(6x^2y + xe^y + \sin y\right) dy = 3x^2y^2 + xe^y - \cos y + h(x).$$

Comparing the two expressions, we can conclude that $f(x,y) = 3x^2y^2 + xe^y - \cos y + C$, for some constant $C$.

This is straightforward enough, but it isn't really integration; it's just *antidifferentiation*. These are not the same thing! They just happen to be related to each other - in single-variable calculus - by the Fundamental Theorem of Calculus. If we want to generalize the concept of *integration*, we have to go back to the original motivating problem, which was that of finding the area below a curve. The corresponding problem for a function of two variables is the problem of finding the *volume below a surface*, and so this is where we'll start.

We'd like to repeat our steps from Math 117, but this time the domain of our function $f(x, y)$ will be a region in the $xy$-plane. For the time being we'll assume this domain to be rectangular: $(a, b) \times (c, d)$ (we'll return to non-rectangular domains soon), and we'll assume that $f > 0$ everywhere in this domain.

- First, partition the $x$-axis into $m$ intervals of length $\Delta x = \frac{(b-a)}{m}$, and partition the $y$-axis into $n$ intervals of length $\Delta y = \frac{(d-c)}{n}$. The effect is to partition the entire domain into rectangles of dimensions $\Delta x$, $\Delta y$ (these rectangles then have area $\Delta A = \Delta x \Delta y$).

- In each of these rectangles, choose a point $(x_i^*, y_j^*)$. The quantity $f(x_i^*, y_j^*)\Delta A$ then represents the volume of a box, which is approximately equal to the volume below the surface $z = f(x, y)$ over this rectangle. See Figure 17.

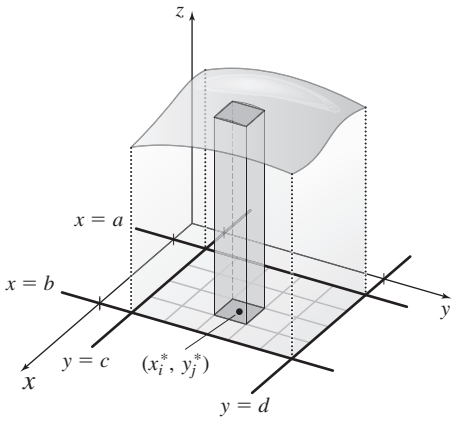



Figure 17:

- Adding the volumes of all of these boxes (there are $mn$ of them) gives an approximation of the volume below the surface over the domain. Furthermore, we expect that the error in the approximation should approach zero as $\Delta A \to 0$.

Based on this, we define the double integral of a continuous function $f(x, y)$ over a rectangular region $R \in \mathbb{R}^2$ as follows:

$$\int_R f(x, y)\, dA = \lim_{m,n\to\infty} \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i^*, y_j^*)\, \Delta A \tag{15}$$

**Notation:** For reasons that we are about to discuss, the double integral is often written $\iint_R f(x, y)\, dA$, and of course we may write $dA$ as $dxdy$.

Note that the definition says nothing about volume, and does not require that $f$ be positive; it is simply a limit of sums of values of $f$. For double integrals over non-rectangular regions, the definition needs to be modified.

### Evaluation of Double Integrals over Rectangles

If $R = [a, b] \times [c, d]$ (a rectangle), then we can evaluate the integral by partial integration, one variable at a time. Here's why:

Suppose we begin by thinking of $y$ as a constant. Then $f(x, y)$ is a function of $x$, and we can integrate it from $x = a$ to $x = b$. Geometrically, *if $f$ is positive*, this amounts to examining a cross-section of the surface $z = f(x, y)$ along the plane $y =$ (constant), and calculating the area below this curve (see Figure 18).

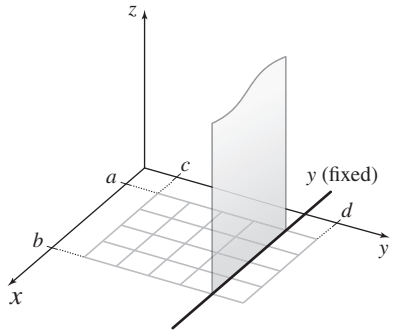



Figure 18:

We now have that the area of a cross-section of the solid whose volume we are trying to compute is $\displaystyle\int_a^b f(x, y)\, dx$. Note that the expression we've obtained has a different value for each value of $y$ we choose; in other words it is a *function* of $y$.

Now, if we multiply our expression by a thickness $\Delta y$, we obtain an approximation for the volume of a "slice" of our solid. Adding up all of these volumes from $y = c$ to $y = d$, and then letting $\Delta y \to 0$, we arrive at the integral *of an integral*! The volume of the solid is

$$\int_c^d \left[ \int_a^b f(x, y)\, dx \right] dy.$$

We call this an *iterated integral*. We normally omit the brackets, but this does not change the meaning. This gives us a method for evaluating double integrals when the domain of

integration is rectangular (and this will be valid even when $f$ is not positive):

$$\text{If } R = [a,b] \times [c,d], \text{ then } \int_R f(x,y)\, dA = \int_c^d \int_a^b f(x,y)\, dxdy. \tag{16}$$

Of course, we could started by thinking of $x$ as constant instead of $y$. Geometrically, this just means starting with a cross-section in the other orientation; see Figure 19.

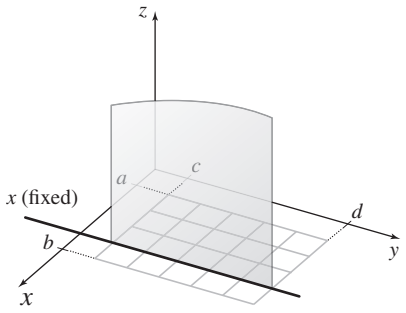



Figure 19:

The resulting area integral is a function of $x$, which we then integrate to find the volume. More generally, we can state an alternative to Equation (16):

$$\text{If } R = [a,b] \times [c,d], \text{ then } \int_R f(x,y)\, dA = \int_a^b \int_c^d f(x,y)\, dydx. \tag{17}$$

The two orders of integration will give the same result, so we have a choice of methods for evaluation of integrals of this type.

**Example:** Evaluate $\displaystyle\int_1^2 \int_0^3 (x^2y^2 + x)\, dxdy$.

**Solution:** First, look at $\displaystyle\int_0^3 (x^2y^2 + x)\, dx$, with $y$ constant:

$$\int_0^3 (x^2y^2 + x)\, dx = \left(\frac{x^3y^2}{3} + \frac{x^2}{2}\right)\Bigg|_0^3 = \frac{27y^2}{3} + \frac{9}{2} = 9y^2 + \frac{9}{2}.$$

Now, integrate this from $y = 1$ to $y = 2$:

$$\int_1^2 \int_0^3 (x^2y^2 + x)\, dxdy = \int_1^2 \left(9y^2 + \frac{9}{2}\right)\, dy = \left(3y^3 + \frac{9y}{2}\right)\Bigg|_1^2 = \ldots = \frac{51}{2}.$$

Alternatively, we could have calculated this as

$$\int_0^3 \int_1^2 (x^2y^2 + x)\, dydx = \int_0^3 \left(\frac{x^2y^3}{3} + xy\right)\Big|_1^2 dx$$

$$= \int_0^3 \left(\frac{7x^3}{9} + \frac{x^2}{2}\right) dx = \ldots = \frac{51}{2}.$$

You'll notice that the intermediate steps look quite different, even though we obtain the same result. This is actually a good thing; it means that if we discover that the integral is difficult (or impossible) to evaluate with one order of integration, we can try the other order of integration instead, and it may work out more easily.

**Example:** Evaluate $\int_0^1 \int_0^{\ln 2} xe^{xy}\, dxdy$.

**Solution:** Notice that as given, the first integral requires integration by parts. However, if we integrate on $y$ first instead, we can avoid this:

$$\int_0^1 \int_0^{\ln 2} xe^{xy}\, dxdy = \int_0^{\ln 2} \int_0^1 xe^{xy}\, dydx$$

$$= \int_0^{\ln 2} e^{xy}|_0^1\, dx$$

$$= \int_0^{\ln 2} (e^x - 1)\, dx$$

$$= (e^x - x)|_0^{\ln 2} = (2 - \ln 2) - (1 - 0) = 1 - \ln 2.$$

Before we move on to a discussion of non-rectangular domains, we mention one special property of integrals over rectangles; if the integrand can be factored (into a function of $x$ and a function of $y$), then the entire integral can be factored (into an integral in $x$ and and integral in $y$). This is easy to prove:

Suppose $f(x, y) = g(x)h(y)$. Then

$$\int_c^d \int_a^b f(x, y)\, dxdy = \int_c^d \int_a^b g(x)h(y)\, dxdy$$

$$= \int_c^d h(y) \int_a^b g(x)\, dxdy.$$

Here we have factored $h(y)$ out of the inner integral, since it's constant with respect to $x$. But now the entire integral $\int_a^b g(x)dx$ is a constant, so it can be factored out of the outer integral to obtain $\int_c^d h(y)\,dy \int_a^b g(x)\,dx$.

**Example:**

$$\begin{aligned}\int_0^1 \int_0^1 \frac{1+x^2}{1+y^2}\,dxdy &= \int_0^1 (1+x^2)dx \int_0^1 \frac{1}{1+y^2}dy \\ &= \left(x+\frac{x^3}{3}\right)\Big|_0^1 \tan^{-1} y\Big|_0^1 \\ &= \left(1+\frac{1}{3}\right)\left(\frac{\pi}{4}\right) = \frac{\pi}{3}\end{aligned}$$

It is never *necessary* to factor an integral in this way, but it may make it a bit easier to avoid errors.

## Evaluation of Double Integrals over Non-rectangular Domains

As we've said, for non-rectangular domains, our definition (15) needs to be modified. We won't actually have to use these new definitions for calculations, but understanding the concepts they make precise is critical. Therefore we'll give outlines of the definitions, without full rigour.

We define a *Type I* region as one which can be described as lying between the graphs of two functions of $x$ over some interval, as shown in Figure 20.

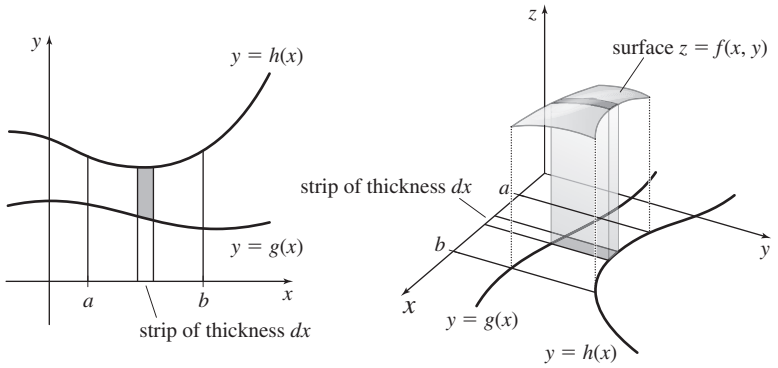



Figure 20:

To calculute the integral of a function $z = f(x, y)$ over such a domain, we start by dividing the domain into thin vertical strips of thickness $\Delta x$ (these strips are approximately rectangular). On each one of these strips we pick one value of $x$, which makes $f$ a function of $y$. We can then integrate this function of $y$ from $y = g(x)$ to $y = h(x)$ to get the area of a cross-section

of the solid, just as before! We can repeat this for every one of our intervals, add the results, and then let $\Delta x \to 0$. That is, we may write

$$\int_R f(x,y)\, dA = \int_a^b \int_{g(x)}^{h(x)} f(x,y)\, dydx. \tag{18}$$

The only difference between this and Equation 17 is that the inner limits of integration are functions of $x$. However, this is an important difference; for one thing it means that we *cannot* change the order of integration so easily anymore[13] !

A *Type II* region is one which can be described as lying between the graphs of two functions of $y$, on some interval, as shown in Figure 21.

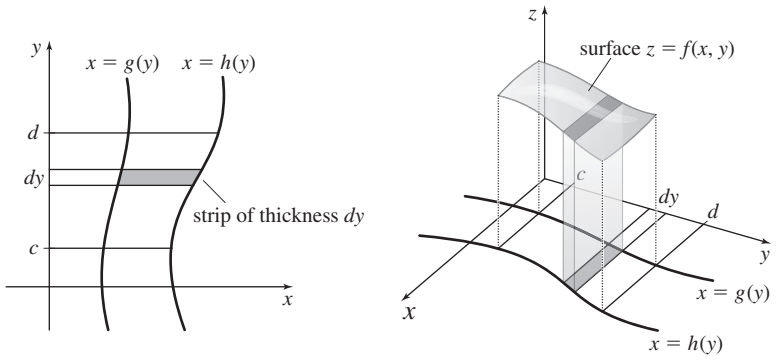



Figure 21:

This time we start by dividing the domain into thin *horizontal* strips of thickness $\Delta y$. On each one of these strips we pick one value of $y$, which makes $f$ a function of $x$. We can then integrate this function of $x$ from $x = g(y)$ to $x = h(y)$, repeat this for every one of our intervals, add the results, and then let $\Delta y \to 0$. The result is

$$\int_R f(x,y)\, dA = \int_c^d \int_{g(y)}^{h(y)} f(x,y)\, dxdy. \tag{19}$$

If a region is neither of Type I nor Type II, then we can subdivide it into smaller regions which are of one type or the other. However, many of the regions you'll encounter can be described in either way; we say these are of *Type III.*

[13]The expression $\int_{g(x)}^{h(x)} \int_a^b f(x,y)\, dxdy$ is NOT a definite integral, because if we evaluate it as an iterated integral we end up with a function of $x$. Therefore it can't possibly the same thing as we have in Equation (19).

**Example:** Find the volume of the solid lying below the surface $z = xy$ and above the region in the $xy$-plane between the functions $y = x$ and $y = \sqrt{x}$.

**Solution:** In these problems our first step must be to sketch the domain, because we need to decide whether to treat it as Type I or Type II. Once that decision is made we need to determine what the limits of integration should be. See Figure 22.

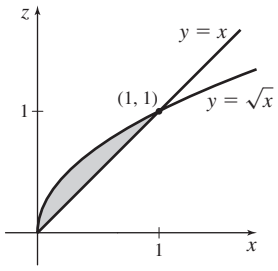



Figure 22:

This particular domain is of Type III; so we can set up the integral in either order:

- First option:

$$\text{Volume} = \iint_R xy\, dA = \int_0^1 \int_x^{\sqrt{x}} xy\, dydx$$

$$= \int_0^1 \left. \frac{xy^2}{2} \right|_x^{\sqrt{x}} dx = \int_0^1 \left( \frac{x^2}{2} - \frac{x^3}{2} \right) dx = \ldots = \frac{1}{24}.$$

- Second option:

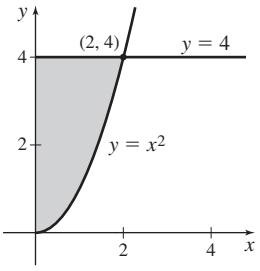
$$\text{Volume} = \iint_R xy\, dA = \int_0^1 \int_{y^2}^{y} xy\, dxdy$$

$$= \int_0^1 \left. \frac{x^2 y}{2} \right|_{y^2}^{y} dy = \int_0^1 \left( \frac{y^3}{2} - \frac{y^5}{2} \right) dy = \ldots = \frac{1}{24}.$$

**Example:** Evaluate $\displaystyle\int_0^2 \int_{x^2}^4 x^3 \sin(y^3)\, dydx$ .

**Solution:** We've been given an integral in iterated form, but we have a problem; we can't do anything with the integral $\displaystyle\int \sin(y^3) dy$. On the other hand, integrating with respect to $x$ would be easy, so it looks as though we should try changing the order of integration. To do this, we'll have to figure out what the domain of integration looks like, which means working backwards from the given limits. We can see that the domain of integration lies between $x = 0$ and $x = 2$, and that for each value of $x$ in this interval the $y$-values run from $y = x^2$ to $y = 4$. Therefore the domain is as shown in Figure 23 (and it is indeed of Type III).

Figure 23:




Describing it the other way, it lies between $y = 0$ and $y = 4$, and for each value of $y$ in this interval the $x$-values run from $x = 0$ to $x = \sqrt{y}$, so we try evaluating our integral this way:

$$\int_0^2 \int_{x^2}^4 x^3 \sin(y^3)\, dydx = \int_0^4 \int_0^{\sqrt{y}} x^3 \sin(y^3)\, dxdy$$

$$= \int_0^4 \left. \frac{x^4}{4} \sin(y^3) \right|_0^{\sqrt{y}} dy$$

$$= \int_0^4 \frac{y^2 \sin(y^3)}{4}\, dy$$

$$= \left. -\frac{\cos(y^3)}{12} \right|_0^4$$

$$= \frac{1}{12}(1 - \cos 64)$$

# 9 Evaluation of Double Integrals in Polar Coordinates

For ordinary single integrals, the most important tool available is the method of substitution. For double integrals $\iint_R f(x,y)\,dA$, we can do something similar; we can change the coordinate system (that is, we change *both* variables!). We'll start with a special case. Polar coordinates may be useful when the domain $R$ is circular, or has circles forming part of its boundary. They may also help if the integrand contains the expression $x^2+y^2$ (since this becomes simply $r^2$). If both of those conditions are met, then changing to polar coordinates is almost definitely the right choice.

So, how do we proceed? Rewriting the integrand is easy enough; we simply set

$$x = \rho\cos\phi \qquad \text{and} \qquad y = \rho\sin\phi$$

and then $f(x,y) = f(\rho\cos\phi, \rho\sin\phi)$. A tougher question is what to do with the "area element" $dA$. In our original development of the double integral this was the area of a rectangle; it was $dxdy$, and so this needs to be modified.

If we partition our range of values of $\rho$ into segments of equal length $d\rho$, and our range of values of $\phi$ into small (equal) angles $d\phi$, then the effect is to subdivide $R$ into small "polar rectangles"... which are not really rectangles at all[14] (see Figure 24).

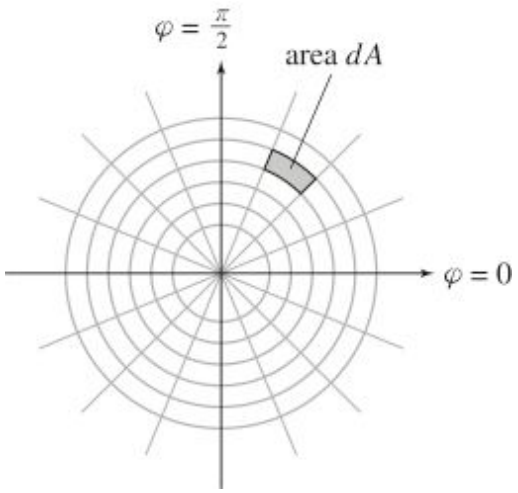



Figure 24:

We need to know what the area of a typical polar rectangle is, so let's take a closer look at one of them (see Figure 25).

[14] To be perfectly correct, we should be writing $\Delta\rho$ and $\Delta\phi$ at this stage, and taking limits as these approach zero as our last step. Think of the current discussion as an outline of the proof of the result to follow.

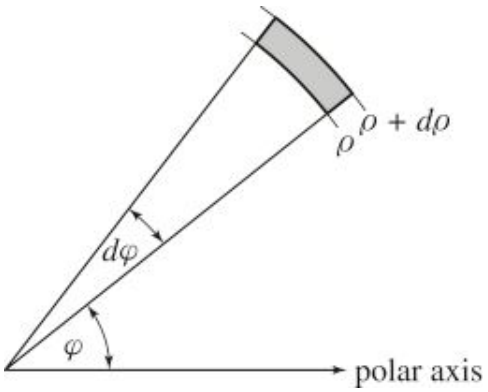


Figure 25:

Recall that for a circle of radius $r$, a sector of angle $\theta$ has area $\frac{1}{2}r^2\theta$. Using this, we can calculate the area of the shaded polar rectangle as the difference between the area of the full sector of radius $\rho + d\rho$ and the smaller (entirely unshaded) sector of radius $\rho$:

$$dA = \frac{1}{2}(\rho + d\rho)^2\, d\phi - \frac{1}{2}\rho^2 d\phi$$

$$= \rho d\rho d\phi + \frac{1}{2}d\rho^2 d\phi.$$

Now, since we're going to be using this in an integral, in which $d\rho$ and $d\phi$ are infinitesimally small, we can discard the second term: (if we write the expression as $dA = d\rho d\phi\left[\rho + \frac{1}{2}d\rho\right]$ we can see that the second term is negligible in comparison to the first). Therefore,

$$\boxed{\iint_{R_{xy}} f(x,y)\, dxdy = \iint_{R_{\rho\phi}} f(\rho\cos\phi, \rho\sin\phi)\, \rho d\rho d\phi} \tag{20}$$

where we've written $R_{xy}$ and $R_{\rho\phi}$ to denote the same domain, described in the appropriate coordinate system.

**Comment:** Now that we have the result, there is a second, more intuitive way to understand it. As $d\rho$ and $d\phi$ approach zero, our typical polar rectangle becomes more and more rectangular. One side of it has length $d\rho$, while the other side has (arc)length $\rho d\phi$, so of course we must have $dA = \rho d\rho d\phi$. Also, notice that the rectangles are not all of the same size; they increase in size with increasing $\rho$, which should make sense looking at Figure 24.

**Example 1:** Evaluate $\iint_R \left(x^3 + xy^2\right) dA$, where $R$ is the portion of the unit circle in the first quadrant.

**Solution:** First notice that the domain is of Type III, so we don't actually *need* polar coordinates; we could express the integral as

$$\iint_R \left(x^3 + xy^2\right) dA = \int_0^1 \int_0^{\sqrt{1-x^2}} \left(x^3 + xy^2\right)\ dydx \tag{21}$$

(or as a similar expression in the other order of iteration). However, the domain is certainly easier to describe in polar coordinates; it's a polar rectangle, in fact. The values of $\phi$ range from 0 to $\frac{\pi}{2}$, and for each value of $\phi$ the values of $\rho$ range from 0 to 1 (see Figure 26).

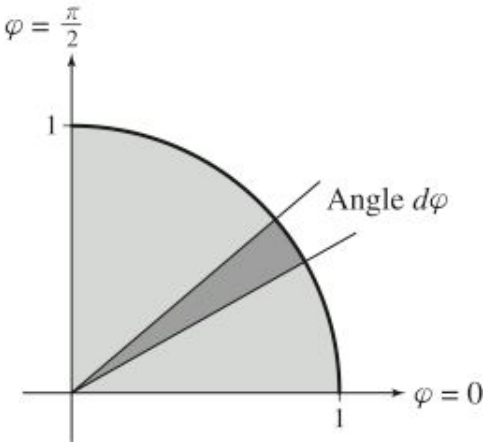



Figure 26:

Also notice that $x^3 + xy^2 = x(x^2 + y^2)$, so this problem satisfies both of the criteria we mentioned at the start. We can now write

$$\iint_R \left(x^3 + xy^2\right) dA = \iint_R x\left(x^2 + y^2\right)\ dA = \int_0^{\frac{\pi}{2}} \int_0^1 (\rho\cos\phi)(\rho^2)(\rho d\rho d\phi)$$

$$= \int_0^{\frac{\pi}{2}} \cos\phi\, d\phi \int_0^1 \rho^4\, d\rho$$

$$= \ldots = \frac{1}{5}.$$

This is certainly an easier calculation than required for the expression in (21).

# 10 The Change-of-Variable Formula

### Background Discussion: Transformations

The name "polar rectangle" is actually more accurate than it might first seem. For a polar rectangle we have $\rho \in [\rho_1, \rho_2]$ and $\phi \in [\phi_1, \phi_2]$, which means that if we sketch this in a Cartesian plane with $\rho$ and $\phi$ as the axes, we see a true rectangle! This is precisely why conversion to polar coordinates was so useful in the last example; in polar coordinates we obtained an integral over a rectangular domain, and that's why we were able to factor it into two separate single integrals. In fact it is possible to view the pair of equations

$$x = \rho\cos\phi \qquad\qquad y = \rho\sin\phi$$

as a *transformation* which maps circles and radial half-lines in the $xy$-plane to horizontal and vertical lines in the first quadrant of the $\rho\phi$-plane (or vice versa). See Figure 27 .

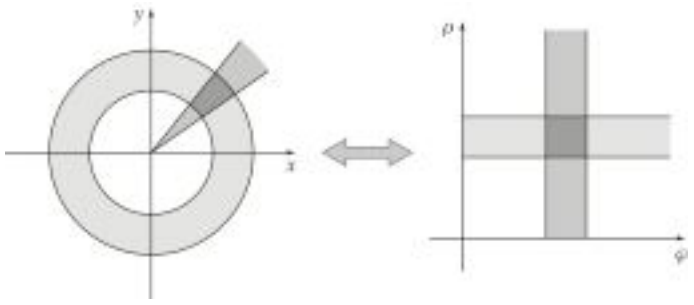


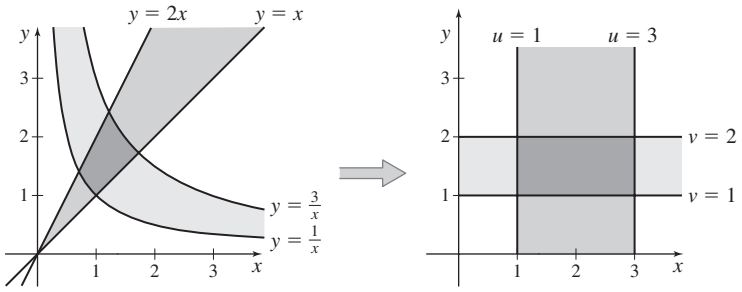
Figure 27:

It is possible to achieve similar results with other transformations. For example, consider the region in the first quadrant bounded by the curves $y = \frac{5}{x}$, $y = \frac{10}{x}$ and the lines $y = x$, $y = 2x$. These equations can be written as $xy = 5$, $xy = 10$, $\frac{y}{x} = 1$ and $\frac{y}{x} = 2$, so if we introduce new variables

$$u = xy, \qquad\qquad v = \frac{y}{x}$$

then these boundaries map to the vertical and horizontal lines $u = 5$, $u = 10$, $v = 1$, and $v = 2$ (see Figure 28).

Figure 28:




In this way we can hope to transform integrals over non-rectangular domains into integrals over rectangular domains (or, at the very least, into integrals whose domains are Type I or Type II regions). The biggest difficulty is that just as for the transformation to polar coordinates, we'll need to determine what the area element $dA$ should be in the new coordinate system. The derivation of the required formula is not all that easy, and so we'll go straight to the result:

**The Change-of-Variable Formula**

Suppose that the variables $x$ and $y$ are related to the variables $u$ and $v$ by the equations $x = x(u, v)$, $y = y(u, v)$. Then

$$\iint_{R_{xy}} f(x,y)\, dxdy = \iint_{R_{uv}} f\left[x\left(u,v\right), y\left(u,v\right)\right] \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \, dudv \tag{22}$$

where

$$\frac{\partial(x,y)}{\partial(u,v)} = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}.$$

This function is called the *Jacobian* of the transformation, and is also denoted by $J$. Note that $\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$ is the *absolute value* of the determinant of the matrix of partial derivatives.

**Comments:**

1. There is one important restriction on the transformation $(x, y) \to (u, v)$; we must not have $J = 0$ at any point on the interior of $R_{uv}$. This condition ensures that the transformation is invertible on the domain of integration. Another way of explaining this is that it ensures that areas in the $xy$-plane get mapped to areas in the $uv$-plane (and not to lines or points).

2. This formula is indeed consistent with Equation 20: if $x = \rho\cos\phi$, and $y = \rho\sin\phi$ then

$$J = \begin{vmatrix} x_\rho & x_\phi \\ y_\rho & y_\phi \end{vmatrix} = \begin{vmatrix} \cos\phi & -\rho\sin\phi \\ \sin\phi & \rho\cos\phi \end{vmatrix} = \rho\cos^2\phi + \rho\sin^2\phi = \rho.$$

   You might notice that in Example 1 we actually do have $\rho = 0$ on the boundary of the domain of integration, but it doesn't happen on the interior, so it doesn't violate the condition in the previous comment. When we move from the Cartesian system to the polar system the origin gets stretched out into the entire line $\rho = 0$, but this is ok! It's the opposite situation that is problematic; if an entire line gets contracted to a point, then we lose all knowledge of the function values along the original line.

3. This is also consistent with the method of substitution of single-variable calculus, although some explanation is required. If we recognize an integral to be of the form $\int_a^b f(g(x))g'(x)dx$, we let $u = g(x)$, and then the integral becomes $\int_{g(a)}^{g(b)} f(u)du$. That is, if $g$ is invertible,

$$\int_{g(a)}^{g(b)} f(u)du = \int_a^b f(g(x))\,\frac{du}{dx}dx.$$

   The details look a little bit different (we seem to be going in the wrong direction), but the key point to notice is that *we always differentiate our substitions* (the derivative $\frac{du}{dx}$ is the single variable version of the Jacobian). For double integrals there are four derivatives instead of one, and the Jacobian combines the information from all four. You might wonder, though, why we have absolute values around the Jacobian for double integrals, but not for single integrals. The reason is that in the single-variable case the sign of the Jacobian $\frac{du}{dx}$ tells us whether the interval of integration has been "flipped around" or not; that is, if $a < b$ and the Jacobian is *negative*, then we'll have $g(a) > g(b)$. When we're evaluating single integrals it doesn't really matter whether the limits of integration go from lowest to highest or not, but *for the double integrals we will always be setting up the limits of integration in the standard order.*

4. The notation $\dfrac{\partial(x,y)}{\partial(u,v)}$ is a reminder of the cancellation that occurs in the single variable case: $\dfrac{dx}{du}du = dx$, and $\dfrac{\partial(x,y)}{\partial(u,v)}dudv = dxdy$.

Perhaps the best way to understand the Jacobian is as a factor which must be introduced to

compensate for the distortion of the domain that occurs when we move it from one coordinate system to another. The following example is intended to illustrate this:

**Example:** Evaluate $I = \iint_D x^2\, dA$, where $D$ is the interior of the ellipse $9x^2 + 4y^2 = 36$.

**Solution:** You may not have evaluated an integral over an ellipse before, but we have just discussed how to evaluate integrals over circles. In fact we can map this ellipse onto a unit circle quite easily; we simply let $u = \frac{x}{2}$ and $v = \frac{y}{3}$ (this means $x = 2u$ and $y = 3v$ , so $9x^2 + 4y^2 = 36 \Longrightarrow 9(4u^2) + 4(9v^2) = 36 \Longrightarrow u^2 + v^2 = 1$). The original ellipse has area $6\pi$, while the new circle has area $\pi$, so this transformation shrinks the domain considerably (see Figure 29).

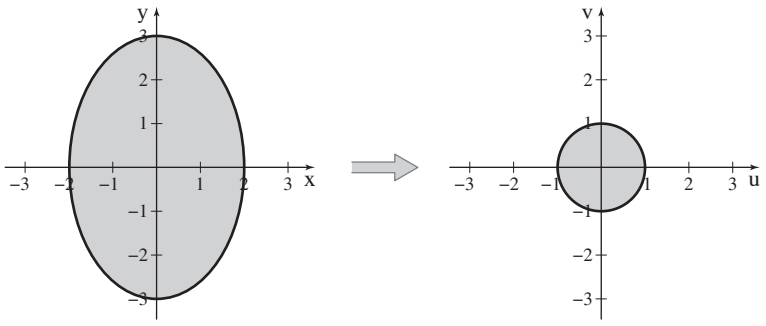



Figure 29:

By formula (22), we have

$$I = \iint_{9x^2+4y^2\le 36} x^2\, dA = \iint_{u^2+v^2\le 1} 4u^2 \frac{\partial(x,y)}{\partial(u,v)} du dv.$$

Now, the Jacobian is $\dfrac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 3 \end{vmatrix} = 6$. The incorporation of this factor into the integral compensates exactly for the shrinking of the domain of integration! We now have $I = \iint_{u^2+v^2\le 1} 24u^2\, dudv$; since the domain is now circular we can evaluate it in polar coordinates. We let $u = \rho\cos\phi$ and $v = \rho\sin\phi$, and then

$$I = \int_0^{2\pi}\int_0^1 (24\rho^2\cos^2\phi)(\rho d\rho d\phi)$$

$$= 24\int_0^{2\pi}\cos^2\phi d\phi \int_0^1 \rho^3 d\rho$$

$$= 24(\pi)\left(\frac{1}{4}\right) = 6\pi$$

**Comment:** We could have done both steps at once, by letting $x = 2\rho\cos\phi$ and $y = 3\rho\sin\phi$. However, since we already know how to work with polar coordinates, it's should be easier to use that knowledge.

**Example:** Evaluate $\iint_D \cos\left(\frac{y-x}{y+x}\right)\, dA$, where $D$ is the trapezoid highlighted in Figure 30.

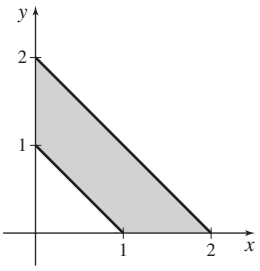


Figure 30:

**Solution:** This problem has, admittedly, been "cooked up" to work out nicely, but it serves to illustrate the method. The boundaries of the domain of integration are $x = 0$, $y = 0$, $y = 1 - x$, and $y = 2 - x$. The latter two can be expressed as $x + y = 1$ and $x + y = 2$, which suggests that $u = x + y$ might be a useful definition for a new variable. Looking at the integrand, we see that the expression $x + y$ appears there as well, and in fact it looks as though the second variable should be defined as $v = y - x$. So, let

$$u = y + x \qquad \text{and} \qquad v = y - x.$$

This transformation will map the boundary $y = 1 - x$ to the line $u = 1$, and it will map the boundary $y = 2 - x$ to the line $u = 2$, but what happens to the other boundaries?

- On the line $x = 0$ we have $u = y$ and $v = y$, so this boundary maps to the line $v = u$.
- On the line $y = 0$ we have $u = x$ and $v = -x$, so this boundary maps to the line $v = -u$.

The image of the transformation is shown in Figure 31;

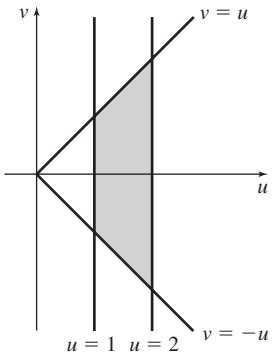


Figure 31:

notice that it is of Type I (we'll think of subdividing it into vertical strips of thickness $du$, stacked side-by-side from $u = 1$ to $u = 2$, and for each value of $u$ in this interval we'll integrate along each strip from $v = -u$ up to $v = u$). Now, we're going to need the Jacobian $\dfrac{\partial(x,y)}{\partial(u,v)}$, but we've given $u$ and $v$ as functions of $x$ and $y$, rather than the other way around. We could invert the transformation[15], but there is an alternative. It can be shown that, as one might hope, $\dfrac{\partial(x,y)}{\partial(u,v)} = \dfrac{1}{\frac{\partial(u,v)}{\partial(x,y)}}$, so we just need to calculate

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2,$$

and this tells us that $\dfrac{\partial(x,y)}{\partial(u,v)} = \dfrac{1}{2}$. We're now ready to finish up:

$$\begin{aligned} \iint_{D_{xy}} \cos\left(\frac{y-x}{y+x}\right)\, dxdy &= \iint_{D_{uv}} \cos\left(\frac{v}{u}\right) \left|\frac{1}{2}\right|\, dudv \\ &= \frac{1}{2}\int_1^2 \int_{-u}^{u} \cos\left(\frac{v}{u}\right)\, dvdu \\ &= \frac{1}{2}\int_1^2 u\sin\left(\frac{v}{u}\right)\Big|_{-u}^{u}\, du \\ &= \frac{1}{2}\int_1^2 \left[u\sin(1) - u\sin(-1)\right] du \\ &= \sin 1 \int_1^2 u du = \sin 1 \left[\frac{u^2}{2}\right]_1^2 = \frac{3}{2}\sin 1. \end{aligned}$$

[15]Not hard to do in this example. Since the transformation is linear, we can add and subtract the equations to and from each other to find $x = \frac{1}{2}(u - v)$, and $y = \frac{1}{2}(u + v)$.

# 11 Discussion: Interpretation of Integrals

So far we have usually interpreted single integrals as areas below curves, and double integrals as volumes below surfaces, although in both cases we have been careful to point out that the *definitions* of these integrals say nothing whatsoever about areas or volumes. These geometric interpretations are useful for understanding the properties of integrals, but in order to be able to use integration in applications you will need to be able to think about them differently. Furthermore, we'll soon be discussing *triple* integrals, and it's very difficult to understand these in terms of geometry![16]

Consider once again the definition of the definite integral:

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*)\Delta x, \tag{23}$$

where $x_i^* \in [x_{i-1}, x_i]$ and $[a, b]$ is divided into $n$ subintervals of equal width $\Delta x$.

This is a limit of a sum of the quantities $f(x_i^*)\Delta x$. The interpretation, therefore, depends on how we interpret $f$ and $x$. **IF** we interpret them both as distances (specifically, if we interpret $\Delta x$ and $f(x_i^*)$ as being the width and height of a thin rectangle), **THEN** the result is an area! In applications, though, the variables will usually have other meanings.

**Example:** If $t$ measures time, and $f(t)$ measures velocity in a straight line, then $\int_a^b f(t)dt$ should be thought of as a sum of infinitesimal *displacements* $ds = f(t)dt$. Therefore it represents the total displacement of a particle moving with velocity $f(t)$ between times $a$ and $b$. Looking at the units of the quantities involved will help: a velocity $f(t)$ might have units of m/s, while $t$ (and $dt$) would have units of s, so $f\,(t)\,dt$ (and $\int_a^b f(t)dt$) must then have units of m.

**Example:** If $\rho(x)$ is the (linear) charge density in a rod of length $L$, with $x$ measuring distance from one end of the rod, then $\int_0^L \rho(x)dx$ must have units of (charge density)$\times$(distance)$=$ $\frac{\text{coulombs}}{\text{meter}}$meters $=$ coulombs; it gives the total charge in the rod.

The same logic applies to double integrals; pay attention to the units!

[16] If you can only think of a single integral as an area below a curve, and you can only think of a double integral as a volume below a surface, then you will be forced to think about a triple integral as a *hypervolume* "below" a three dimensional object, in four dimensions. This is probably *not* something you want to think too hard about!

**Example:** If $f(x,y)$ is a population density (in organisms/km$^2$), then the expression $f\,(x,y)\,dA$ has units of (organisms/km$^2$)$\times$(km$^2$)=(organisms), so $\iint_R\ f(x,y)\ dA$ gives the total population within the domain $R$.

### Integrals with no Integrand

Here's a special case that will seem peculiar at first: an integral doesn't actually have to have an integrand at all!

- The integral $\int_a^b dx$ is a sum of infinitesimal lengths $dx$, and we recognize it as giving the length of the interval of integration: $\int_a^b dx = b - a$.
- The integral $\iint_R dA$ is a sum of infinitesimal areas $dA$. It gives the area of the domain $R$!

Believe it or not, this can actually be a useful tool for finding areas, if it happens that $R$ can be described conveniently in a different coordinate system. Here's an example to convince you that it actually works:

**Example:** Find the area of the region illustrated in Figure 32.

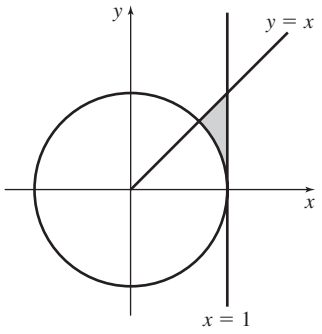



Figure 32:

**Solution:** First, we note that this problem can be solved without calculus; we can calculate the area of the shaded region as the difference between the area of the triangle and the area of the sector of the circle: $A = \frac{1}{2} - \frac{\pi}{8}$. We want to show that using the double integral $\iint_R dA$ yields the same result. There are two obvious ways to evaluate this:

- In Cartesian coordinates, $\iint_R dA = \int_{\frac{1}{\sqrt{2}}}^{1} \int_{\sqrt{1-x^2}}^{x}\ dydx\ =\ \int_{\frac{1}{\sqrt{2}}}^{1} \left(x - \sqrt{1-x^2}\right) dx$. This *can* be evaluated, but for the second term we'd need a trigonometric substitution; we'll

find another way instead. Note, however, that we could have written down exactly the same integral with our Math 117 techniques. In fact $\iint_R dA$ will *always* reduce to $\int_a^b [f_{upper}(x) - f_{lower}(x)]\,dx$ or $\int_c^d [g_{right}(y) - g_{left}(y)]\,dy$ if we remain in Cartesian coordinates. The advantage to writing $\iint_R dA$ is that it gives us the option of choosing a different coordinate system, as we're about to do!

- In Polar coordinates, the circle has equation $\rho = 1$, while the straight line $x = 1$ has equation $\rho\cos\phi = 1$, i.e. $\rho = \sec\phi$. Therefore, realizing that $\phi$ runs from 0 to $\frac{\pi}{4}$, and that for each value of $\phi$ in this interval $\rho$ runs from $\rho = 0$ to $\rho = \sec\phi$, we have

$$\iint_R dA = \int_0^{\frac{\pi}{4}} \int_1^{\sec\phi} \rho d\rho d\phi = \int_0^{\frac{\pi}{4}} \left.\frac{\rho^2}{2}\right|_1^{\sec\phi} d\phi$$

$$= \frac{1}{2}\int_0^{\frac{\pi}{4}} \left(\sec^2\phi - 1\right) d\phi$$

$$= \left.\frac{1}{2}\left(\tan\phi - \phi\right)\right|_0^{\frac{\pi}{4}}$$

$$= \frac{1}{2} - \frac{\pi}{8}.$$

You'll see another example or two in the assignments.

**Mean Values**

Here's one other special application of the definite integral; it can be used to calculate average values of functions over specified domains. We discussed the single-variable case in Math 117, and showed that the average value of a function $f(x)$ over an interval $[a, b]$ can be calculated as

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

Similarly, the average value of a function $f(x, y)$ over a two-dimensional region $R$ can be calculated as

$$f_{avg} = \frac{1}{\text{Area}(R)}\iint_R f(x, y)dA.$$

**Example:** Find the average distance from the origin of all the points within the unit circle.

**Solution:** The distance of a point $(x, y)$ from the origin is given by $f(x, y) = \sqrt{x^2 + y^2}$, and so

$$\begin{aligned} Average\ Distance &= \frac{\iint_R \sqrt{x^2 + y^2} dA}{\iint_R dA} \\ &= \frac{\int_0^{2\pi} \int_0^1 \rho \left(\rho d\rho d\phi\right)}{\pi} \\ &= \frac{1}{\pi} \int_0^{2\pi} d\phi \int_0^1 \rho^2 d\rho \\ &= \left(\frac{1}{\pi}\right)(2\pi)\left(\frac{\rho^3}{3}\Big|_0^1\right) \\ &= \frac{2}{3}. \end{aligned}$$

**All of these ideas can now be extended to functions of *three* variables!**

- $\iiint_D dV$ gives the volume of the (three-dimensional) region $D$.
- $\iiint_D f(x, y, z) dV$ should be interpreted as a sum of the infinitesimal quantities $f\ dV$ over all of the points in $D$.
- $\dfrac{1}{\text{Volume}(D)} \iiint_D f(x, y, z) dV$ gives the mean value of the function $f$ on the region $D$.

## 12 Triple Integrals

If you've understood the preceding discussion, then the concept of a triple integral should be a natural extension of the concept of (single) integrals and double integrals. We will now be speaking of functions of three variables, and our domains of integration will be three-dimensional. As we've done for double integrals, we'll avoid rigorous definitions, and instead give outlines which will (we hope) enable you to construct the integrals you need.

If the domain of integration (call it $D$) is a rectangular box, with $x \in [a_1, a_2]$, $y \in [b_1, b_2]$, $z \in [c_1, c_2]$, then the jump from two variables to three is an easy one. We partition each of the three axes, into intervals of lengths $\Delta x$, $\Delta y$, and $\Delta z$ respectively. In each of the resulting small boxes we choose a point $(x_i^*, y_j^*, z_k^*)$, evaluate the function (let's call it $f$) at that point, and multiply the result by the volume of the box $\Delta V = \Delta x \Delta y \Delta z$ (see Figure 33).

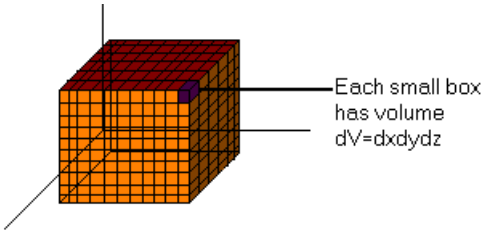



Figure 33:

We then sum the results for every one of these boxes within the domain. The limit of the result as $\Delta x$, $\Delta y$, $\Delta z$ all approach zero is our integral, $\int_D f(x,y,z)\, dV$ (and just as we often use a double integral sign when integrating functions of two variables, we can use a triple integral sign now, if we wish, and write $\iiint_D f(x,y,z)\, dV$). Of course, there are now *six* possible orders of integration, and for rectangular domains these can all be interchanged freely:

$$\begin{aligned}\int_D f(x,y,z)\, dV &= \int_{c_1}^{c_2}\int_{b_1}^{b_2}\int_{a_1}^{a_2} f(x,y,z)\, dxdydz \\ &= \int_{c_1}^{c_2}\int_{a_1}^{a_2}\int_{b_1}^{b_2} f(x,y,z)\, dydxdz \\ &= \int_{b_1}^{b_2}\int_{c_1}^{c_2}\int_{a_1}^{a_2} f(x,y,z)\, dxdzdy \\ &\text{etc.}\end{aligned}$$

For domains which are *not* rectangular, we can extend the concept of "Type I", "Type II", and "Type III" regions... but we can now distinguish between at least *seven* types. Numbering them isn't necessary; if we grasp the concept we should be able to write down an appropriate

integral.

**Example 1:** Suppose the domain of integration can be described[17] as $D = \{\ (x, y, z) \mid a \leq x \leq b,\ \phi_1(x) \leq y \leq \phi_2(x),\ \gamma_1(x, y) \leq z \leq \gamma_2(x, y)\ \}$, as shown in Figure 34 (this is the three-dimensional "Type I" region). This isn't easy to show with a static 2-D image, but what we're describing is a solid whose front and back sides are flat (they are in the planes $x = a$ and $x = b$), and whose left and right sides are curved in one dimension only, meaning that you could easily mold a sheet of paper to them with no crumpling. The top and bottom, meanwhile, may curve in two dimensions (they are unrestricted *surfaces*).

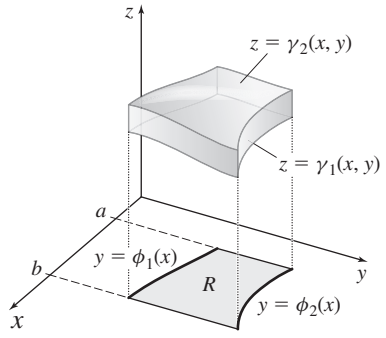


Figure 34:

The integral of a function $f$ over such a domain can be expressed as

$$\int_D f(x, y, z)\, dV = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} \int_{\gamma_1(x,y)}^{\gamma_2(x,y)} f(x, y, z)\, dz dy dx. \tag{24}$$

To understand why it must have this form, consider that a *definite integral* is a number, so the last pair of limits of integration we use must be constants. That is, the outer integral must be the integral in $x$, from $x = a$ to $x = b$. Now, for each value of $x$ in the interval $[a, b]$, the values of $y$ run from $y = \phi_1(x)$ to $y = \phi_2(x)$. These limits are constant with respect to the remaining variables $y$ and $z$, so we can make this our middle integral. Finally, for each point $(x, y)$ in the region labelled $R$, the values of $z$ in the domain $D$ run from $z = \gamma_1(x, y)$ up to $z = \gamma_2(x, y)$.

It might help to view a triple integral as a double integral *of* a single integral[18]. Viewed this way, Equation (24) is a double integral over $R$ of the function $g(x, y) = \displaystyle\int_{\gamma_1(x,y)}^{\gamma_2(x,y)} f(x, y, z)\, dz$.

[17] Please don't let the notation confuse you here; we've used $\phi$ as a function name here, just because it's the Greek counterpart to $f$. It is *not* being used as an angle here!

[18] You can also view it as a single integral of a double integral, if you prefer.

Setting up a double integral should be familiar now; $R$ is a Type I (2-D) region, so we write $\int_a^b \int_{\phi_1(x)}^{\phi_2(x)} g(x,y)\, dydx$. Of course, for this to be helpful we need to understand what the function $g$ represents: it can be thought of as the sum of all the values of $f$ along a vertical column, one of which extends above each point $(x,y)$ in the region $R$ from $z = \gamma_1(x,y)$ up to $z = \gamma_2(x,y)$. See Figure 35.

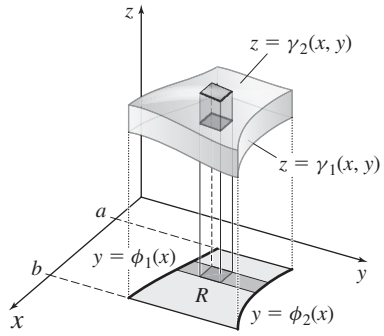



Figure 35:

**Example 2:** Evaluate $\iiint_D xz\, dV$, where $D$ is the region in the 1st octant (where $x$, $y$, and $z$ are all positive) below the plane $z = y$ and inside the cylinder (see Figure 36; the second image is a view from a different perspective, with the extra sections of the plane and cylinder removed).

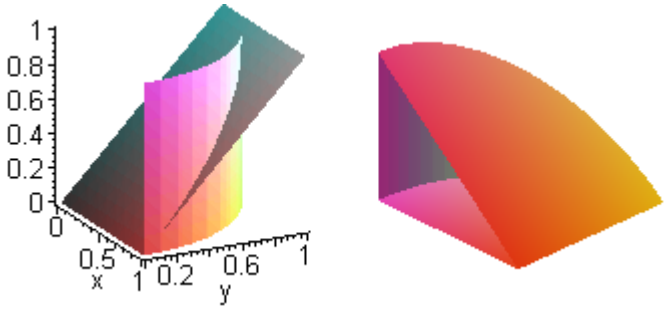



Figure 36:

**Solution:** This happens to be a Type I region. Viewed from above, the entire domain lies within the quarter circle shown in Figure 37.

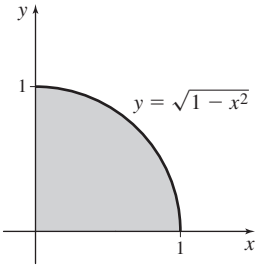



Figure 37:

It is bounded below by the plane $z = 0$ and above by the plane $z = y$, so we can set the integral up as

$$\iiint_D xz\, dV = \int_0^1 \int_0^{\sqrt{1-x^2}} \int_0^y xz\, dzdydx = \ldots = \frac{1}{30}. \tag{25}$$

We'll leave the evaluation of this expression as an exercise (you should get 1/30).

### Cylindrical Coordinates

Looking at the previous example, it might occur to you that the exterior double integral could be more conveniently expressed in polar coordinates. In fact this is exactly what we'll do, but it's traditional to deal with the triple integral as a whole, and for this we need a three-dimensional change of coordinates. The *cylindrical* coordinate system is a trivial extension of the polar coordinate system; a point in $\mathbb{R}^3$ can be described as in Figure 38

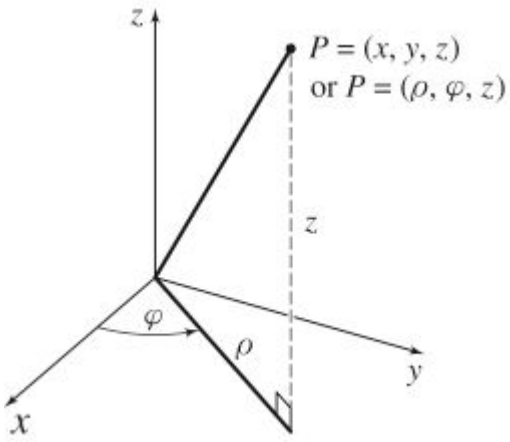



Figure 38:

. The transformation equations are simply these:

$$x = \rho \cos \phi$$

$$y = \rho \sin \phi \tag{26}$$

$$z = z$$

The form of the Jacobian is easily extended to three variables, and the result for this particular

case will not be a surprise:

$$\frac{\partial(x,y,z)}{\partial(\rho,\phi,z)} = \begin{vmatrix} x_\rho & x_\phi & x_z \\ y_\rho & y_\phi & y_z \\ z_\rho & z_\phi & z_z \end{vmatrix} = \begin{vmatrix} \cos\phi & -\rho\sin\phi & 0 \\ \sin\phi & \rho\cos\phi & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho.$$

Therefore, when switching from Cartesian to cylindrical coordinates, we can always replace $dz\ dy\ dx$ with $\rho\ dz\ d\rho\ d\phi$ (and in those cases where we wish to use cylindrical coordinates, this will almost always be the desired order of the differentials - you'll almost never want to write $\rho\ d\rho\ d\phi\ dz$ or any of the other possibilities).

**Example 3:** Completing the example we started above, we can calculate

$$\begin{aligned} \iiint_D xz\ dV &= \int_0^1 \int_0^{\sqrt{1-x^2}} \int_0^y xz\ dzdydx \\ &= \int_0^{\frac{\pi}{2}} \int_0^1 \int_0^{\rho\sin\phi} (\rho\cos\phi)(z)(\rho dz d\rho d\phi) \\ &= \int_0^{\frac{\pi}{2}} \int_0^1 \int_0^{\rho\sin\phi} \rho^2 z\cos\phi dz d\rho d\phi \\ &= \int_0^{\frac{\pi}{2}} \int_0^1 \rho^2 \frac{z^2}{2}\cos\phi \Big|_{z=0}^{z=\rho\sin\phi} d\rho d\phi \\ &= \int_0^{\frac{\pi}{2}} \int_0^1 \frac{\rho^4\sin^2\phi\cos\phi}{2} d\rho d\phi \\ &= \frac{1}{2}\int_0^{\frac{\pi}{2}} \sin^2\phi\cos\phi d\phi \int_0^1 \rho^4 d\rho \\ &= \frac{1}{2}\left(\frac{\sin^3\phi}{3}\Big|_0^{\frac{\pi}{2}}\right)\left(\frac{\rho^5}{5}\Big|_0^1\right) = \frac{1}{2}\left(\frac{1}{3}\right)\left(\frac{1}{5}\right) = \frac{1}{30}. \end{aligned}$$

Notice that we replaced both $x$ in the integrand and $y$ in the inner limit of integration with the appropriate cylindrical expressions.

**Spherical Coordinates**

There is a second way to generalize the concept of polar coordinates to three dimensions, which is simultaneously more natural and more complicated. Just as the polar coordinate system in $\mathbb{R}^2$ use the distance from the origin and one angle, the *spherical* coordinate system in $\mathbb{R}^3$ uses the distance from the origin and *two* angles. The distance $r$ is measured directly from the origin to the point (in $\mathbb{R}^3$), so $r = \sqrt{x^2 + y^2 + z^2},\ r \in [0, \infty)$. The angle $\phi$ is defined exactly the same way as the angle $\phi$ we use for cylindrical coordinates, so $\phi \in [0, 2\pi)$. Meanwhile, the angle $\theta$ is the angle away from the positive $z$-axis; for this we need only $\theta \in [0, \pi]$ (since the farthest we can get away from the positive $z$-axis is the negative $z$-axis)[19]. See Figure 39 (we've included the polar coordinates $\rho$ and $\phi$ for reasons you'll see in a moment).

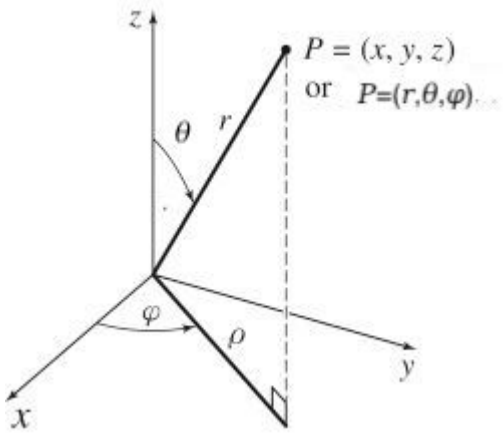



Figure 39:

Actually, spherical coordinates won't be a completely new concept for you; the idea is almost identical to the usage of latitude and longitude to identify locations on the surface of the earth. The angle $\phi$ corresponds exactly to the measurement of longitude (with the Greenwich meridian defining the $xz$-plane, as 0°), while the angle $\theta$ differs from the measurement of latitude only in that the reference point is the "north pole" instead of the equator; if we measured latitude using $\theta$ then the values would range from 0° at the north pole to 180° at the south pole, instead of from 90°N to 90°S. The other difference is a simple one; the coordinate $r$ allows us to specify locations which don't lie exactly *on* the earth's surface, but above or below it.

[19]We have the same problem with notation for cylindrical and spherical coordinates that we have for polar coordinates. In these notes, we have used the convention which is common in physics, to try to be consistent with your other courses. However, in most mathematical texts (including your online text) you'll see $r$ and $\rho$ swapped from what we've done here, as well as $\theta$ and $\phi$.

For the conversion formulas, notice from Figure 39 that we can relate the spherical coordinates easily to the cylindrical coordinates: $z = r\cos\theta$ and $\rho = r\sin\theta$. This allows us to make the connection to Cartesian coordinates:

$$\begin{aligned} x &= \rho\cos\phi = r\sin\theta\cos\phi \\ y &= \rho\sin\phi = r\sin\theta\sin\phi \\ z &= \quad z \quad = \quad r\cos\theta \end{aligned} \tag{27}$$

Of course, we'll need the Jacobian of this transformation:

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{vmatrix} x_r & x_\theta & x_\phi \\ y_r & y_\theta & y_\phi \\ z_r & z_\theta & z_\phi \end{vmatrix}$$

$$= \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}$$

$$= \ldots = r^2\sin\theta.$$

We'll let you fill in the missing lines of the calculation on your own. We can use the result as a formula: whenever we express a triple integral in spherical coordinates we will express $dV$ as $r^2\sin\theta\ dr\ d\theta\ d\phi$.

**Example 4:** Recalling the discussion of the previous section, we should be able to calculate the volume of a sphere of radius $a$ as $\iiint_D dV$. Now, if $D$ is a sphere, then spherical coordinates are the most sensible tool. To describe the entire sphere, we need the full range for both angles. The volume is

$$V = \iiint_{D_{xyz}} dV = \iiint_{D_{r\phi\theta}} r^2\sin\theta\ dr\ d\phi\ d\theta$$

$$= \int_0^\pi \int_0^{2\pi} \int_0^a r^2\sin\theta\ dr\ d\phi\ d\theta$$

$$= \int_0^{\pi} \sin\theta \, d\theta \int_0^{2\pi} d\phi \int_0^{a} r^2 \, dr$$

$$= (2)(2\pi)\left(\frac{a^3}{3}\right)$$

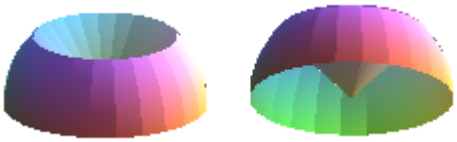
$$= \frac{4}{3}\pi a^3,$$

which is the result we should have been expecting.

**Example 5:** Find the average value of $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ over the region lying above the $xy$-plane, inside the sphere $x^2 + y^2 + z^2 = 4$, and below the cone $z = \sqrt{x^2 + y^2}$.

**Solution:** The domain is illustrated (from two different perspectives) in Figure 40; we'll call it $E$.

Figure 40:



We are looking for

$$f_{\text{avg}} = \frac{\iiint_E f \, dV}{\text{Volume}(E)} = \frac{\iiint_E f \, dV}{\iiint_E dV}.$$

In spherical coordinates, the integrand $\sqrt{x^2 + y^2 + z^2}$ is simply $r$, and $dV = r^2 \sin\theta \, dr \, d\theta \, d\phi$. We also need to describe the boundaries of the domain in this system: the sphere has the simple equation $r = 2$, but what about the cone? Well,

$$z = \sqrt{x^2 + y^2} \qquad \text{means} \qquad r\cos\theta = \sqrt{r^2 \sin^2\theta \cos^2\phi + r^2 \sin^2\theta \sin^2\phi}$$

$$= r \sin\theta,$$

so $\tan\theta = 1$, meaning that $\theta = \pi/4$. Upon reflection this *should* have been obvious; any cone will have equation $\theta = \kappa$, for some constant $\kappa$, and our cone opens up at an angle of 45° from the $z$-axis. If you think about that for a moment, it should also become clear that the spherical equation of our third boundary, the $xy$-plane, is $\theta = \pi/2$. Putting all of this information to

use, then, we have

$$\iiint_E \sqrt{x^2+y^2+z^2}\, dV = \int_0^{2\pi} \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_0^2 r \cdot r^2 \sin\theta \; dr \; d\theta \; d\phi$$

$$= \int_0^{2\pi} d\phi \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \sin\theta \; d\theta \int_0^2 r^3 \; dr$$

$$= (2\pi)\left(\frac{\sqrt{2}}{2}\right)(4) = 4\sqrt{2}\pi.$$

Meanwhile,

$$\iiint_E dV = \int_0^{2\pi} \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_0^2 r^2 \sin\theta \; dr \; d\theta \; d\phi$$

$$= \int_0^{2\pi} d\phi \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \sin\theta \; d\theta \int_0^2 r^2 \; dr$$

$$= (2\pi)\left(\frac{\sqrt{2}}{2}\right)\left(\frac{8}{3}\right) = \frac{8\sqrt{2}\pi}{3}.$$

Therefore the mean value of $\sqrt{x^2+y^2+z^2}$ within $E$ is $\dfrac{4\sqrt{2}\pi}{8\sqrt{2}\pi/3} = \dfrac{3}{2}$.

Suggestion: think about what the stated problem was here, and convince yourself that our result is realistic. Then go back and do the same for Example 2.

## Part II

# Taylor Polynomials and Series

## 13 Introduction

In Math 117 you were introduced to the basic tools of calculus, and you have already begun to see that there are some problems for which we simply cannot find exact solutions. As an example, recall that the function $\sin\left(x^2\right)$ has no antiderivative in terms of the functions that we commonly use. Unfortunately, these problems are not unusual, and so we often find that the only option available to us is to seek an *approximate* solution.

We have two basic strategies for finding these approximate solutions: numerical methods and analytical methods. Numerical methods generally use our mathematical definitions by "brute force". For example, to approximate the value of $\displaystyle\int_0^{0.2} \sin\left(\pi x^2\right) dx$ we could refer to the definition of the definite integral, and calculate the area of $n$ rectangles of width $\dfrac{0.2}{n}$ and heights determined by the function $\sin\left(\pi x^2\right)$. Analytical methods, on the other hand, use the theory of calculus (also known as *analysis*) to recognize reasonable approximations for the functions involved. As an example, we could argue that if $x$ is small, then $\sin x \approx x$. Therefore $\sin\left(\pi x^2\right) \approx \pi x^2$ when $\pi x^2$ is small, and so $\displaystyle\int_0^{0.2} \sin\left(\pi x^2\right) dx \approx \int_0^{0.2} \pi x^2 dx$, which we can evaluate easily (we get $\pi/375$, which is correct to within 0.2%).

Each of the two approaches has its own strengths and weaknesses. With the advances in computer technology over the last several decades the numerical approach has become extremely powerful, but analytical methods still have their place:

- Analytical methods allow us to find approximations for functions, without specifying the numerical values involved. In a numerical approach we must usually assign variables to all of the parameters involved, whether we know what they should be or not.

- Anaylytical methods can often be used to check that our numerical methods are giving realistic results. This is helpful, because mistakes can easily go unnoticed when you're asking a computer to do most of the work!

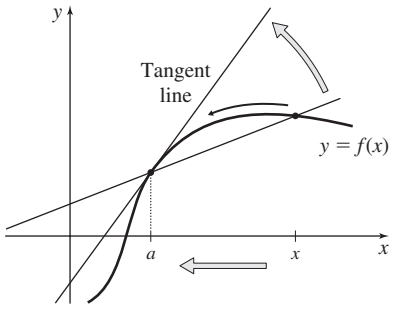
In these notes you will find a brief introduction to some common numerical methods, but the course will concentrate on the analytic approach. Specifically, we'll be looking at how

polynomials can be used as approximations for more complicated functions.

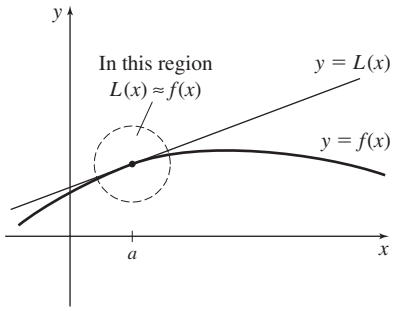
## 14 Our Simplest Option: the Linear Approximation (a.k.a. the Tangent Line Approximation, or Linearization)

We reviewed the concept of linear approximations earlier in this course, but it may be worth reviewing again briefly, since it's a foundational concept for what is to come:

The derivative $f'(a)$ is defined as $\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$. Graphically, we interpret this as the slope of the tangent line to $y = f(x)$ at $x = a$:




For values of $x$ **near** $a$, the tangent line gives a reasonable approximation to $f(x)$:




What is the function $L(x)$? Recall: the line through $(a, b)$ with slope $m$ has equation $(y - b) = f'(a)(x - a)$, i.e.

$$L(x) = y = f(a) + f'(a)(x - a).$$

This can be helpful when the function $f(x)$ is easy to evaluate at $x = a$ but difficult to work with at points nearby.

**Example:** Consider the function $f(x) = \sqrt{x}$. We know that $f(4) = 2$, but what is $f(3.98)$? It's hard to evaluate $\sqrt{3.98}$ exactly (without a calculator), but the equation of the tangent line is much simpler:

$$L(x) = f(4) + f'(4)(x-4)$$

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

$$\implies L(3.98) = 2 + \frac{1}{4}(-0.02) = 2 - 0.005 = 1.995.$$

This allows us to conclude that $\sqrt{3.98} \approx 1.995$.

We can perform the same calculation more concisely using the notation of differentials: $f(a + \Delta x) = f(a) + \Delta f$, where $\Delta f \approx f'(a)\,\Delta x$. In our example we identify $a = 4$, $\Delta x = -0.02$, and then calculate $\Delta f \approx \frac{1}{4}(-0.02) = -0.005$, and so $f(3.98) \approx 2 - 0.005 = 1.995$.

**Example:** The resistivity of a wire of a certain metal depends on the temperature according to the equation $\rho(t) = \rho_{20} e^{\alpha(t-20)}$ Ohm-meters, where $\rho_{20}$ is the resisitivity of the metal at $20°C$ and $\alpha$ is another material property of the metal called the temperature coefficient. This formula is often replaced with its linear approximation at $t = 20$:

$$\rho'(t) = \alpha\rho_{20} e^{\alpha(t-20)}, \text{ so } \left.\begin{array}{ll} \rho(20) & = \rho_{20} \\ \rho'(20) & = \alpha\rho_{20} \end{array}\right\} \implies L(t) = \rho_{20} + \alpha\rho_{20}(t-20).$$

This is easier to calculate, and it gives excellent approximations for temperatures much less than $\frac{1}{\alpha}$.

For example, for copper the values of the constants are $\rho_{20} = 1.7 \times 10^{-8}\Omega$-m and $\alpha = 0.0039/°C$ (so $\frac{1}{\alpha} \approx 256°C$). At $40°C$, then,

$$\rho \approx 1.7 \times 10^{-8} + (0.0039)\left(1.7 \times 10^{-8}\right)(20)$$

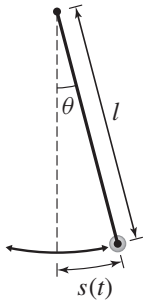
$$= 1.8 \times 10^{-8}\Omega\text{-m},$$

which matches the precise value to our two significant digits.

**Example:** You're already accustomed to the approximation $\sin(x) \approx x$ for values of $x$ near zero. You might not have realized it, but this IS a linearization!

$$\left.\begin{array}{lll} f(x) & = \sin x & f(0) = 0 \\ f'(x) & = \cos x & f'(0) = 1 \end{array}\right\} \Longrightarrow \left\{\begin{array}{ll} L(x) & = 0 + 1(x - 0) \\ & = x \end{array}\right.$$

**Example: The simple pendulum** Consider the motion of a body of mass $m$ suspended by a rod of length $l$ from a fixed point. In developing a mathematical model of a physical problem such as this, we'll need some preliminary approximations before we even get to the mathematics! Here we'll assume (or pretend?) that the mass is concentrated at a point, that the rod is massless and rigid, and that there is no air resistance or friction. That is, we'll consider gravity as the only force in effect.




If we let $s(t)$ be the displacement along the arc, as in the figure, (with $s$ understood to be positive when the angle $\theta$ is positive), then Newton's second law of motion ($F = ma$) tells us that

$$m\frac{d^2s}{dt^2} = -mg\sin\theta.$$

Now, there are too many variables here, but we can eliminate one of them. We know that $s = l\theta$, so $\dfrac{d^2s}{dt^2} = l\dfrac{d^2\theta}{dt^2}$, which means that $ml\dfrac{d^2\theta}{dt^2} = -mg\sin\theta$. That is,

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\sin\theta.$$

If we specify our initial conditions (say $\theta(0) = \theta_0$ and $\theta'(0) = 0$), then our model is complete. Unfortunately, we can't solve it! At this point we need another approximation. We know that if the amplitude of the oscillations remains small (that is, as long as $\theta$ is close to zero), then

the equation above is approximately equivalent to the equation

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\theta,$$

which is so much simpler that we can *guess* the solution for the initial conditions we've specified! We're looking for a function whose second derivative is just $-\frac{g}{l}$ times itself... this must be a combination of cosines and sines of $\sqrt{\frac{g}{l}}t$, and to match our initial conditions the only possibility is $\theta(t) = \cos\left(\sqrt{\frac{g}{l}}t\right)$.

Comment: we're going to explore how we can improve on the linear approximation (using quadratic approximations, cubic approximations, and so on). However, in some examples, such as this pendulum problem, *linearity* turns out to be the key. If we were to replace $\sin\theta$ with a more accurate approximation, we wouldn't be able to solve the problem!

# 15 Root Finding

As we discussed last time, part of the motivation for the second half of this course is the inconvenient fact that we cannot evaluate integrals such as $\int e^{-x^2}dx$ exactly. In fact, though, these intractable integrals aren't the first examples of unsolvable problems that you've encountered; finding roots of algebraic equations can also be impossible![20]

For example, suppose we want to know where $x = e^{-x}$. A quick sketch tells us that there must be such a value of $x$, but we cannot solve for it algebraically. However, we can approximate it using numerical methods, and there are several options available.

## 15.1 The Bisection Method

The simplest approach is to take advantage of the intermediate value theorem (IVT). For example, to find out where $x = e^{-x}$, we could let $f(x) = x - e^{-x}$ (so now we're trying to find out where $f(x) = 0$). Notice that $f(0) = -1$ (so it's negative), while $f(1) \approx 0.63$ (which is positive). Since $f(x)$ is continuous on the interval in between, we can apply the IVT and conclude that there must be a root in that interval.

The idea of the bisection method is that if we bisect the interval $(0, 1)$ into the two intervals

[20] Finding the "root" of an equation $f(x) = g(x)$ means solving for $x$. If the equation is $f(x) = 0$, then we may also speak of finding the "zeroes" of $f$.

$(0, 0.5)$ and $(0.5, 1)$, we can use the IVT again to determine which subinterval contains the root. Checking the value of $f$ at the midpoint, we have $f(0.5) \approx -11$ . Since this is negative, and $f(1)$ is already known to be positive, we know that the root (let's call it $x^*$) lies in the interval $(0.5, 1)$.

We can repeat this indefinitely:

$f(0.75) \approx 0.28 > 0$, so $x^* \in (0.5, 0.75)$

$f(0.625) \approx 0.09 > 0$, so $x^* \in (0.5, 0.625)$

$f(0.5625) \approx -0.007 < 0$, so $x^* \in (0.5625, 0.625)$

Of course, this is very slow, so for now let's conclude for the moment that the root is approximately 0.6 (to one decimal place), and move on.

## 15.2 Newton's Method (aka the Newton-Raphson Procedure)

An alternative method was suggested by Isaac Newton himself. The idea is simple; if we can't solve the equation $f(x) = 0$, then replace $f(x)$ with a linear approximation $L(x)$, and solve the equation $L(x) = 0$ instead!

**Example:** Here's Newton's own original example: find a root of the equation $x^3 - 2x - 5 = 0$. We want to find a linear approximation to the function $f(x) = x^3 - 2x - 5$... but where should it be based? That is, what should the point of tangency, $x_0$, be?

Well, if we want the approximation to be a good one, then we need $x_0$ to be close to the actual root. So, we'll *start* with the Bisection Method!

Consider a few values of $f$:

$f(0) = -5$

$f(1) = -6$

$f(2) = -1$

$f(3) = 16$

... and now we can see that there must be a root in the interval $(2, 3)$.

In fact, it looks as though it's likely to be closer to 2 than to 3, so let's use $x_0 = 2$.

Next, we want the linearization of $f(x)$ at $x_0 = 2$:

$$f(x) = x^3 - 2x - 5 \qquad f(2) = -1$$
$$f'(x) = 3x^2 - 2 \qquad f'(2) = 10$$

$$\implies L_2(x) = -1 + 10(x-2)$$

$$= 10x - 21$$

So, here's Newton's argument: since $x^3 - 2x - 5$ is approximately equal to $10x - 21$ when $x$ is close to 2, we expect that $x^3 - 2x - 5 = 0$ approximately where $10x - 21 = 0$, that is, at $x = 2.1$.

Now, we can repeat this! We expect that 2.1 is a better approximation of the root than 2 was, so let's give it a name: $x_1 = 2.1$, and find the linear approximation to $f(x)$ at $x_1$:

$$f(2.1) = 0.061$$
$$f'(2.1) = 11.23$$

$$\implies L_{2.1}(x) = 0.061 + 11.23(x - 2.1)$$

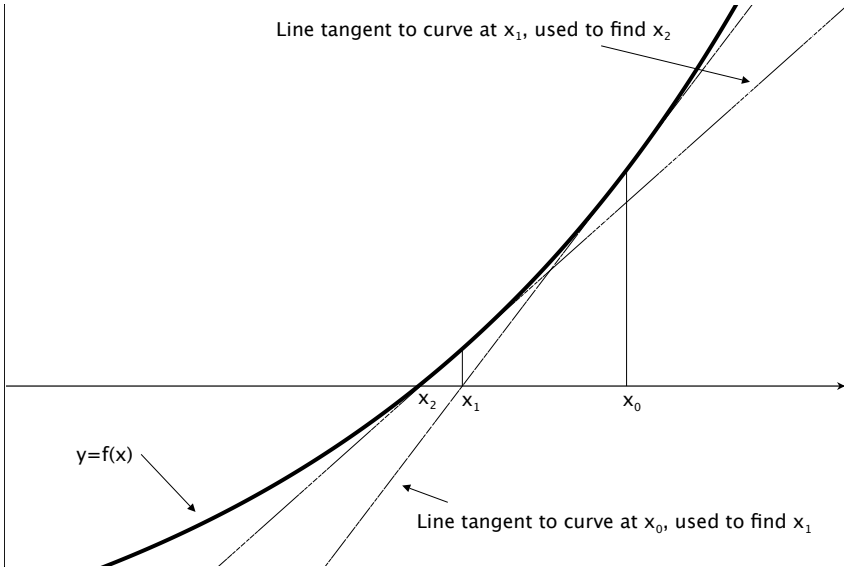
$$= 11.23x - 23.522.$$

Setting $L_{2.1}(x) = 0$ gives $x_2 \approx 2.09457$. We can check that this is working: observe that $f(x_2) \approx 0.0002$.

Continuing in this way, we can generate a sequence of values $x_0, x_1, x_2, \ldots$ which converges to a zero of $f(x)$.

### Ilustration

The figure here is for a different example ($f(x) = e^{x^2} - 2$, with $x_0 = 1$), chosen just for ease of illustration. We make a rough guess of the root, and call that guess $x_0$. We then find out where the tangent line to $f$ at $x_0$ hits the $x$-axis... and you can see that this brings us closer to the actual root. We then label this new point as $x_1$, and find the tangent line to $f$ at $x_1$. Finding out where that line hits the $x$-axis takes us even closer to the root. In fact, it's so close that in the figure it we can no longer see any space at all between $x_2$ and the root!




### The Iterative Formula

Rather than calculating linear approximations step-by-step and case-by-case, as above, we can derive a general formula as follows:

- Pick $x_0$ (using the Bisection Method, or possibly just by using a sketch).
- The linear approximation to $f(x)$ at $x_0$ is

$$L_{x_0}(x) = f(x_0) + f'(x_0)(x - x_0).$$

- Set this equal to zero, solve for $x$, and call the result $x_1$:

$$f(x_0) + f'(x_0)(x - x_0) = 0$$

$$\implies x - x_0 = -\frac{f(x_0)}{f'(x_0)}$$

$$\implies x = x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

- Repeat, using the linear approximation to $f(x)$ at $x_1$. This will give

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

- Repeating indefinitely establishes that in general,

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$

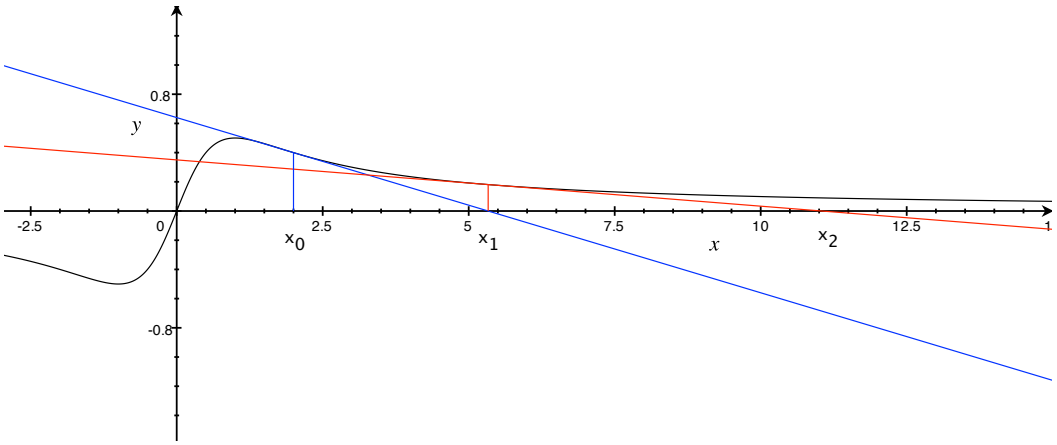
**Notes:**

- The sequence $\{x_n\}$ will not converge if $f'(x)$ fails to exist or isn't continuous at the root. Fortunately, in such cases we're unlikely to need the method in the first place, because we should be able to find these kinds of points using analytical methods! For example, we cannot use Newton's Method to locate the zero of $f(x) = \sqrt[3]{x}$... but why would we bother?

- A similar problem occurs if $f'(x) = 0$ at the root. In this cases the sequence may still converge, but probably not with its usual speed.

- Less obviously, the method may not work well if $f''(x)$ is infinite at the root. As a simple example, consider the function $f(x) = x^{4/3}$. You can easily show that the iterative formula simplifies to $x_{n+1} = \frac{1}{4}x_n$. This sequence will converge, but it's not much quicker than the bisection method!

- Aside from those two possible problems, our sequence should converge rapidly, *as long as our initial guess is good enough.* If we make a poor guess, then our sequence may converge to a root other than the one we want, or it might even diverge entirely. For example, the figure below shows what would happen if we tried the method on the function $f(x) = \dfrac{x}{x^2+1}$, with an initial guess of $x_0 = 2$. This sort of problem really isn't as bad as it sounds... we can always fix it by making a better guess, and it isn't that hard to find one using the Bisection Method.

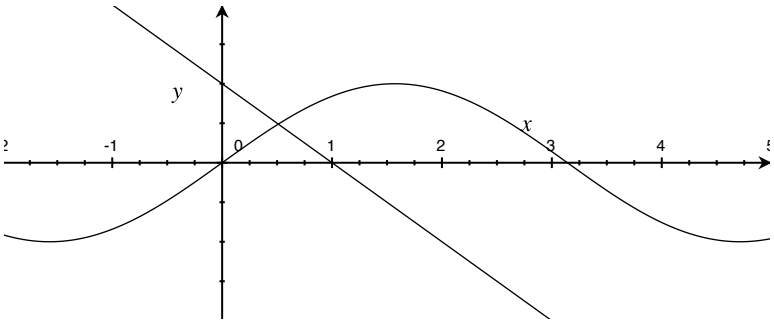



**Example**

Find a root of the equation $\sin x + x - 1 = 0$, correct to six decimal places.

**Solution:**

From a quick sketch it is obvious that there is exactly one root, and that a reasonable first guess would be $x_0 = 0.5$.




Now, letting $f(x) = \sin x + x - 1$, we have $f'(x) = \cos x + 1$, and so

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$= 0.5 - \frac{(-0.0206)}{1.8776}$$

$$= 0.510957953$$

$$\hookrightarrow \quad x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

$$= 0.510973429$$

$$\hookrightarrow \quad x_3 = 0.510973429.$$

At this point, all 9 digits on my calculator have stopped changing, so we can conclude that this is the root, to nine decimal places of accuracy! This is actually quite typical; with a good first guess, the sequence will usually converge *very* quickly.

## 15.3 Fixed-Point Iteration (optional section)

Although Newton's Method usually converges in very few steps, the individual calculations can be somewhat tedious. Fixed-point iteration is an alternative method which typically takes more iterations, but in which the individual steps are easier. The idea is simple: rewrite the equation $f(x) = 0$ as $x = g(x)$, by isolating $x$ in some way (or, if necessary, just by adding $x$ to both sides of the equation)[21]. We can then hope to find an approximation of a root by using the recurrence relation $x_{n+1} = g(x_n)$.

### Example

Let's revisit our bisection method example, and locate the point where $x = e^{-x}$. This is already in the required form; our iteration formula is just

$$x_{n+1} = e^{-x_n}.$$

We concluded earlier that the root was approximately 0.6, so let's use that as $x_0$. We then find that

$$x_1 = e^{-0.6} \approx 0.5488$$

$$x_2 = e^{-0.5488} \approx 0.5776$$

$$x_3 = e^{-0.5776} \approx 0.5612$$

$$x_4 = e^{-0.5612} \approx 0.5705$$

$$x_5 = e^{-0.5705} \approx 0.5652$$

$$x_6 = e^{-0.5652} \approx 0.5682$$

[21] A solution of the equation $x = g(x)$ is called a *fixed point* of $g$.

$$x_7 = e^{-5682} \approx 0.5665$$

$$x_8 = e^{-0.5665} \approx 0.5675$$

$$x_9 = e^{-0.5675} \approx 0.5669$$

$$x_{10} = e^{-5669} \approx 0.5673$$

$$x_{11} = e^{-0.5673} \approx 0.5671$$

$$x_{12} = e^{-0.5671} \approx 0.5672$$

$$x_{13} = e^{-0.5672} \approx 0.5671$$

$$x_{14} = e^{-0.5671} \approx 0.5672$$

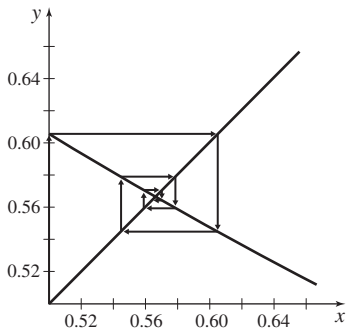
$$x_{15} = e^{-0.5672} \approx 0.5671$$

$$x_{16} = e^{-0.5671} \approx 0.5671$$

and we can conclude that the root is 0.5671, to 4 decimal places. Note: If the last few steps here look a bit odd, it's because I've used more digits in my calculations than I've written down. The root turns out to be very close to 0.56715, which explains the apparent flipping back and forth between 0.5671 and 0.5672.

**Comment:** If you've ever tried hitting the "sin" button on a calculator over and over, and watched the numbers shrink, then you've solved the equation $\sin x = x$ by fixed point iteration!

**Geometric Explanation:** To see why this method (sometimes) works, consider this diagram:



- Picking $x_0 = 0.5$, we evaluate $f = 0.607$, and we can think of travelling to the point $(0.5, 0.607)$.
- Using that output as our next input means setting $x_{n+1} = y_n$, and we can interpret this as taking us horizontally to the line $y = x$. We're now at the point $(0.607, 0.607)$.
- Evaluating $f(x)$ again carries us vertically (down) back to the curve $y = e^{-x}$, and the point $(0.607, 0.545)$.
- As we continue, we repeatedly from line to curve and from curve to line, and we can see that end up following a "square spiral" towards the root.

If you think about this a bit, you'll realize that the reason we spiral inwards is that the vertical distances are always smaller than the preceeding horizontal distances, which occurs here because the slope of $f(x)$ is less than one in magnitude (it's negative, but $> -1$). If, instead, we had a function with $f'(x) < -1$, we would spiral outwards instead.

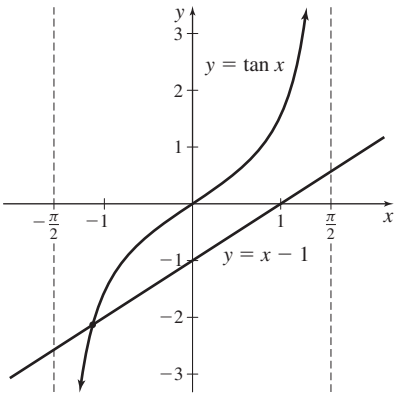
Now, what happens if $f'(x)$ is positive? If you think about this, you'll realize that we will still get convergence if $f'(x) < 1$... although we don't move in a spiral! (What *does* happen?) Therefore we can state one simple requirement:

**Theorem: Convergence of Fixed-Point Iteration** Suppose that $f(x)$ is defined for all $x \in \mathbb{R}$, that it is differentiable everywhere, and that its derivative is always bounded (so that there are no points with vertical tangents). If the equation $f(x) = x$ has a solution (i.e. if $f(x)$ has a fixed point), and **if $|f'(x)| < 1$ for all values of $x$ within some interval containing the fixed point,** then the sequence generated by letting $x_{n+1} = f(x_n)$, will converge, with *any* choice of $x_0$ in that interval.

Of course, we'll usually be able to tell if our sequence is convergent just by trying it, so if it's difficult to tell if the condition $|f'(x)| < 1$ is satisfied, we don't need to worry too much. If our sequence turns out to be divergent, or if it converges to a root other than the one we're looking for, we can often try isolating $x$ in a different way (or using Newton's Method!).

**Example** Locate the solution (in the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$) of the equation $\tan x = x - 1$.

**Solution:** A sketch verifies that there is only one solution:




Now, if we write $x = \tan x + 1$ and use $x_{n+1} = \tan x_n + 1$, we can see that the sequence will *diverge*, because

$$\frac{d}{dx}(\tan x + 1) = \sec^2 x, \qquad \text{and } \left|\sec^2 x\right| = \frac{1}{\cos^2 x} \geq 1 \quad \text{(for all } x\text{)}.$$

You can try it if you wish; the sequence diverges in a hurry!

However, we could also write the problem as $x = \tan^{-1}(x - 1)$. This will *work*, because

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

and this is $\leq 1$ for all $x$. We can therefore use the sequence $x_{n+1} = \tan^{-1}(x_n - 1)$. Any first point will work; let's try $x_0 = 0$:

$$x_1 = \tan^{-1}(-1) \approx -0.7854$$

$$x_2 = \tan^{-1}(-1.7854) \approx -0.1060$$

$$x_3 \approx -1.1889$$

$$\ldots$$

$$x_8 \approx x_9 \approx -1.1323,$$

and that's our root, to four decimal places.

# 16 Polynomial Interpolation

The goal in the lectures to come will be to discuss how we can improve on the idea of the tangent line approximation, but in order to do that we need one more digression. We hope that you'll find this material to be useful in its own right.

Suppose we are given $n+1$ points, which we'll label as $(x_0, y_0)$, $(x_1, y_1)$, $\ldots (x_n, y_n)$, and we would like to find a smooth curve which passes through all of them. The simplest such curve will be a polynomial of degree $n$ (for two points the simplest option is a line, for three points it is a parabola, for four points it is a cubic, and so on).

For simplicity, let's begin by supposing that $n = 3$, and that the $x$ values are

$$x_0 = 0, \quad x_1 = 1, \quad x_2 = 2, \quad x_3 = 3.$$

How do we proceed? Well, we want a cubic, so we write

$$y = a + bx + cx^2 + dx^3. \tag{28}$$

If we "plug in" our four points $(0, y_0)$, $(1, y_1)$, $(2, y_2)$, and $(3, y_3)$, we find

$$\begin{aligned} y_0 &= a \\ y_1 &= a + b + c + d \\ y_2 &= a + 2b + 4c + 8d \\ y_3 &= a + 3b + 9c + 27d. \end{aligned} \tag{29}$$

We now have a system of four equations in four unknowns, for the coefficients $a$, $b$, $c$, and $d$. We *could* just solve this system with our familiar linear algebra techniques. However, Newton, who didn't have standard linear algebra techniques at his disposal, found a useful formula for the interpolating polynomial by solving for the coefficients in an ingenious way. Here's what he did:

First, he introduced some new notation. We'll define the *first finite differences* to be

$$\begin{aligned}\Delta y_0 &= y_1 - y_0\\ \Delta y_1 &= y_2 - y_1\\ \Delta y_2 &= y_3 - y_2\end{aligned}$$

(and so on, if we have more than four points).

Calculating these quantities from the original system(29), we find

$$\begin{aligned}\Delta y_0 &= b + c + d\\ \Delta y_1 &= b + 3c + 7d\\ \Delta y_2 &= b + 5c + 19d,\end{aligned}$$

and you'll notice that all of the $a$'s have disappeared; we've reduced the problem to three equations in three unknowns!

Next, we continue by defining the *second finite differences* as $\Delta^2 y_n = \Delta\left(\Delta y_n\right) = \Delta y_{n+1} - \Delta y_n$, which gives

$$\begin{aligned}\Delta^2 y_0 &= 2c + 6d\\ \Delta^2 y_1 &= 2c + 12d.\end{aligned}$$

Finally, the *third finite difference*, defined generally by $\Delta^3 y_n = \Delta^2 y_{n+1} - \Delta^2 y_n$, is

$$\Delta^3 y_0 = 6d.$$

Now we can solve for our coefficients (in terms of the finite differences):

$$\begin{aligned}d &= \tfrac{1}{6}\Delta^3 y_0\\ c &= \tfrac{1}{2}\left(\Delta^2 y_0 - \Delta^3 y_0\right)\\ b &= \Delta y_0 - \tfrac{1}{2}\Delta^2 y_0 + \tfrac{1}{3}\Delta^3 y_0\\ a &= y_0.\end{aligned}$$

Notice that we've kept only the finite differences involving the zero subscript. Plugging these results into (28) and collecting terms (by order of finite difference) gives us (after a little bit of work)

$$y = y_0 + x\Delta y_0 + x\left(x-1\right)\frac{\Delta^2 y_0}{2} + x\left(x-1\right)\left(x-2\right)\frac{\Delta^3 y_0}{6}.$$

### First Generalization

You can probably spot the pattern in the formula above, and guess that if we repeat these calculations using *more* than four points we'll obtain this:

$$\boxed{\begin{array}{c} y = y_0 + x\Delta y_0 + x\,(x-1)\,\frac{\Delta^2 y_0}{2!} + x\,(x-1)\,(x-2)\,\frac{\Delta^3 y_0}{3!} + \ldots \\ \ldots + x\,(x-1)\,(x-2)\cdots(x-n+1)\,\frac{\Delta^n y_0}{n!} \end{array}} \tag{30}$$

This will look much simpler once you realize that we can calculate the finite differences very easily. If we're working by hand, we'll usually use a triangular table; we start by writing down all the $y$-values, then write the first finite differences beside them, and then the second finite differences beside those, and so on... like this:

$$\begin{array}{lllll} y_0 & & & & \\ & \Delta y_0 & & & \\ y_1 & & \Delta^2 y_0 & & \\ & \Delta y_1 & & \Delta^3 y_0 & \\ y_2 & & \Delta^2 y_1 & & \Delta^4 y_0 \\ & \Delta y_2 & & \Delta^3 y_1 & \\ y_3 & & \Delta^2 y_2 & & \\ & \Delta y_3 & & & \\ y_4 & & & & \end{array}$$

All the coefficients we need end up in the top diagonal row.

**Example** Find the 4th-order polynomial which passes through the points $(0, 2)$, $(1, 2)$, $(2, 12)$, $(3, 62)$, and $(4, 206)$.

**Solution:** Find the finite differences:

$$\begin{array}{ccccc}
2 & & & & \\
& 0 & & & \\
2 & & 10 & & \\
& 10 & & 30 & \\
12 & & 40 & & 24 \\
& 50 & & 54 & \\
62 & & 94 & & \\
& 144 & & & \\
206 & & & &
\end{array}$$

Inserting them into formula (30) yields the following expression:

$$\begin{aligned}
y &= y_0 + x\Delta y_0 + \tfrac{1}{2!}x\,(x-1)\,\Delta^2 y_0 + \tfrac{1}{3!}x\,(x-1)\,(x-2)\,\Delta^3 y_0 \\
&\quad + \tfrac{1}{4!}x\,(x-1)\,(x-2)\,(x-3)\,\Delta^4 y_0 \\
&= 2 + 0 + 5x\,(x-1) + 5x\,(x-1)\,(x-2) + x\,(x-1)\,(x-2)\,(x-3)\,.
\end{aligned}$$

We won't normally bother simplifying these expressions, but this one happens to simplify to $y = 2 - x + x^2 - x^3 + x^4$.

### Generalization to Non-integer Nodes

Let's now remove the restriction that the "nodes" $x_0$, $x_1$, $x_2$, .... be precisely the numbers 0, 1, 2, etc. We will still require them to be equidistant, however. That is, we'll require that $x_n = x_0 + nh$ for $n = 0, 1, 2, \ldots$ (so $h$ is the distance between each pair of nodes). Non-equidistant nodes can also be dealt with, but we won't investigate that case in this course.

To find the appropriate form of the interpolating polynomials, we would repeat the procedure we used to find formula (30). The calculations are a bit messier, so we won't go through them here; instead we'll just state the result.

**The Newton Forward Difference Formula** Given $n+1$ equidistant nodes $x_0$, $x_n = x_0 + nh$, the $n^{\text{th}}$-order polynomial passing through all of them is given by

$$y = y_0 + \frac{(x - x_0)}{h}\Delta y_0 + \frac{(x - x_0)(x - x_1)}{2!h^2}\Delta^2 y_0 + \dots$$
$$\dots + \frac{(x - x_0)(x - x_1)\cdots(x - x_{n-1})}{n!h^n}\Delta^n y_0 \tag{31}$$

You should be able to see immediately that if we set $x_0 = 0$ and $h = 1$, then we recover our previous formula.

**Example:** Estimate the value of $f(2.45)$ if $f(x)$ passes through the points $(2, 4)$, $(2.2, 5)$, $(2.4, 4)$, and $(2.6, 2)$.

**Solution:** The finite differences are

$$\begin{array}{ccccc} 4 & & & & \\ & 1 & & & \\ 5 & & -2 & & \\ & -1 & & 1 & \\ 4 & & -1 & & \\ & -2 & & & \\ 2 & & & & \end{array}$$

and so we obtain the cubic

$$y = 4 + \frac{(x-2)}{0.2}(1) + \frac{(x-2)(x-2.2)}{2!(0.2)^2}(-2) + \frac{(x-2)(x-2.2)(x-2.4)}{3!(0.2)^3}(1)$$

$$= 4 + 5(x-2) - 25(x-2)(x-2.2) + \frac{125}{6}(x-2)(x-2.2)(x-2.4).$$

Evaluating this at 2.45 suggests that $f(2.45) \approx 3.555$.

**Comments:**

- Notice that the last node, 2.6, doesn't appear explicitly in our formula. However, the information about $f$ at this node *was* included in the calculation of $\Delta^3 y_0$.
- Keep in mind that we have *absolutely no knowledge* about the error in our approximation, because we have no information about how $f$ behaves between the nodes. We are simply

assuming that its behaviour is predictable, and finding the simplest polynomial that matches the information we *do* have.

### Linear Interpolation

If you try a few examples, you'll discover that polynomial interpolation can give surprising results. If we don't trust the results, it may be more sensible to use just the closest two points for each approximation. Our formula (31) is usually written differently in this case.

First, we write $\Delta y_0$ as $y_1 - y_0$ (which is how we defined it in the first place), and $h$ as $x_1 - x_0$. Then (31) reads

$$y = y_0 + \frac{(x - x_0)}{(x_1 - x_0)} (y_1 - y_0) .$$

This can be re-arranged:

$$y = \left[1 - \frac{(x - x_1)}{(x_1 - x_0)}\right] y_0 + \frac{(x - x_0)}{(x_1 - x_0)} y_1$$

$$= \frac{(x - x_1)}{(x_0 - x_1)} y_0 + \frac{(x - x_0)}{(x_1 - x_0)} y_1 .$$

This is known as the *Lagrange Linear Interpolation Formula.*

**Exercise:** Show that linear interpolation gives $f(2.45) \approx 3.5$ in our previous example.

# 17 Taylor Polynomials

We've now discussed two ideas which are both credited to Isaac Newton:

- For a smooth function $f(x)$, a tangent line to the graph at a point $(x_0, f(x_0))$ can be defined by considering the secant line joining $(x_0, f(x_0))$ to a second point $(x_1, f(x_1))$ and letting $x_1 \to x_0$. This is how we defined the derivative (as the slope of that tangent line), and we can use the tangent line as an approximation to $f(x)$ for values of $x$ near $x_0$.
- For any set of $n+1$ points with equidistant nodes, a polynomial of degree $n$ can be found which passes through all of them.

Some 40 years later, around 1715, it occurred to an Englishman named Brook Taylor that these ideas could be combined: if we have multiple points, what happens to the interpolating polynomial if we allow them *all* to merge?

For simplicity, let's start with three points. Consider a function $f : x \to f(x)$ at the equidistant points

$$x_0 = x_0$$

$$x_1 = x_0 + \Delta x$$

$$x_2 = x_0 + 2\Delta x.$$

The corresponding values of $y$ will be

$$y_0 = f(x_0)$$

$$y_1 = f(x_1) = f(x_0 + \Delta x)$$

$$y_2 = f(x_2) = f(x_0 + 2\Delta x).$$

Applying the Newton Forward Difference Formula, the equation of the parabola joining these three points is

$$y = y_0 + (x - x_0)\frac{\Delta y_0}{\Delta x} + \frac{1}{2}(x - x_0)(x - x_1)\frac{\Delta^2 y_0}{(\Delta x)^2}.$$

So, what happens as $\Delta x$ approaches zero? First of all, we know that $\dfrac{\Delta y_0}{\Delta x} \to f'(x_0)$ as $\Delta x \to 0$

(that's how we defined $f'(x_0)$ in Math 117). That is, in the first two terms we have reproduced the equation of the tangent line, $y = y_0 + f'(x_0)(x - x_0)$. What about the third term?

Taylor assumed[22], correctly, that $\dfrac{\Delta^2 y_0}{(\Delta x)^2}$ would approach $f''(x_0)$ as $\Delta x \to 0$. Recalling also that we're letting $x_1 \to x_0$, we discover that when we let the three points merge into one, we arrive at the expression

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

So, what have we discovered? We have an object which resembles the tangent line, but uses more information: $f(x_0)$, $f'(x_0)$, and $f''(x_0)$, instead of just $f(x_0)$ and $f'(x_0)$. Since it gives the equation of a parabola instead of a line, we call it the *quadratic approximation* to $f(x)$. A little experimentation shows that it is indeed (usually) a more accurate approximation than the linear one:

**Example:** Consider the function $f(x) = \ln x$. To find its linear approximation for $x$ near 1, we need the values of $f$ and $f'$ at 1: $f(1) = 0$, and $f'(x) = 1/x$, so $f'(1) = 1$. Hence the linear approximation is

$$\ln x \approx 0 + 1(x - 1) .$$

That is, $\ln x \approx x - 1$, when $x$ is close to 1.

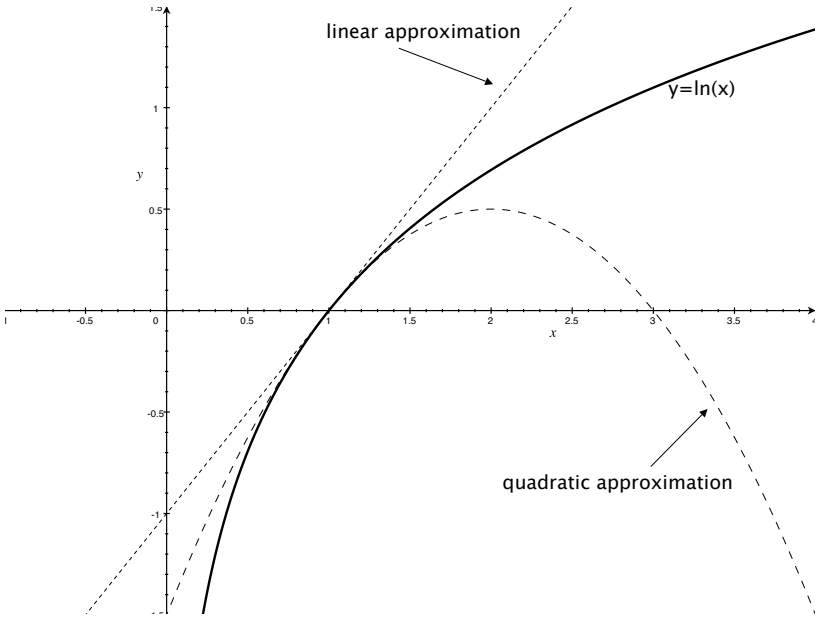
For the quadratic approximation, we calculate also $f''(x) = -1/x^2$, which gives $f''(1) = -1$. The quadratic approximation is therefore

$$\ln x \approx 0 + 1(x - 1) - \frac{1}{2}(x - 1)^2 .$$

That is, $\ln x \approx x - 1 - \dfrac{1}{2}(x - 1)^2$ (when $x$ is close to 1).

The figure below shows the graphs of $f(x)$ and its linear and quadratic approximations; clearly the quadratic approximation is an improvement on the linear one, because it's able to match the *concavity* of the original function, which is something a line can't do!

[22] Taylor couldn't prove this, because the concept of limits hadn't been developed yet! As it turns out, the proof is more difficult than you might expect... it's certainly beyond the level of this course. However, it has indeed been proven, so we'll accept it as fact.




Of course, we don't have to stop here. If we go back and repeat the procedure using *more* than three points, say $n+1$ of them, we obtain a polynomial of degree $n$, which should be an even better approximation to $f$ for values of $x$ near $x_0$ (for our purposes we'll just have to assume that $\dfrac{\Delta^k y_0}{(\Delta x)^k}$ approaches $f^{(k)}(x_0)$ as $\Delta x$ approaches zero - which it does). We call this the *nth-order Taylor polynomial centered at* $x_0$, and give it a special notation:

$$P_{n,x_0}(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2!}f''(x_0)(x-x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n.$$

This can be written more concisely in summation notation as

$$\boxed{P_{n,x_0}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}. \tag{32}$$

(Now is probably a good time to tell you that you will be expected to know this formula!)

**Example:** Find the 3rd-order Taylor polynomial for $f(x) = \ln x$, centered at 1.

**Solution:** We've already calculated the 2nd-order polynomial above, so all we have to do is find the cubic term. Since $f'''(x) = 2/x^3$, we have $f'''(1) = 2$, and so

$$\begin{aligned} P_{3,1}(x) &= 0 + 1(x-1) - \tfrac{1}{2!}(x-1)^2 + \tfrac{2}{3!}(x-1)^3 \\ &= (x-1) - \tfrac{1}{2}(x-1)^2 + \tfrac{1}{3}(x-1)^3. \end{aligned}$$

To get a sense of how these polynomials behave, let's compare their values to the values of $\ln x$ at a few different values of $x$:

| $x$ | $\ln x$ | $P_{1,1}(x)$ | $P_{2,1}(x)$ | $P_{3,1}(x)$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 1.1 | 0.09531... | 0.1 | 0.095 | $0.095\overline{33}$ |
| 1.2 | 0.18233... | 0.2 | 0.18 | $0.18\overline{266}$ |
| 1.3 | 0.26236... | 0.3 | 0.255 | 0.264 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 2 | 0.69314... | 1 | 0.5 | $0.8\overline{3333}$ |
| 3 | 1.09861... | 2 | 0 | $2.\overline{666}$ |

The first few entries here appear to back up our hypothesis; as we add more terms, our approximations become more accurate. We should also not be surprised that the approximations get worse as we move farther away from $x = 1$. However, if you look at the last line, you'll see some very bad news; ***if we stray too far away from*** $x = 1$***, then not only are the approximations not very good, but they actually get*** WORSE ***as we add more terms!***

If if we consider values of $x$ less than 1 we get similar results, but the situation at $x = 0$ is catastrophic:

| $x$ | $\ln x$ | $P_{1,1}(x)$ | $P_{2,1}(x)$ | $P_{3,1}(x)$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 |
| 0.9 | -0.10536... | -0.1 | -0.105 | $-0.105\overline{33}$ |
| 0.8 | -0.22314... | -0.2 | -0.22 | $-0.22\overline{266}$ |
| 0.7 | -0.35667... | -0.3 | -0.345 | $-0.354$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 0 | Undefined | -1 | -1.5 | $-1.8\overline{3333}$ |

This shouldn't really be a surprise; there's no way that any polynomial approximation is going to be able to reproduce a vertical asymptote!

The rest of the first half of this course will be devoted to developing efficient ways to calculate Taylor polynomials, and determining the range of values of $x$ for which we can use them successfully. It turns out that the polynomials in the above example "work" only for the

values of $x$ in the interval $(0, 2]$, and by the time of the midterm exam, you will (we hope) understand why!

## 17.1 Preview of Applications

Why is all this important? It is simply because polynomials are easier to work with than other functions. If we encounter a problem which is too difficult for us to solve, we may be able to find an approximate solution by replacing any inconvenient functions with their polynomial approximations.

**Basic Approximations** As an example, consider the problem of calculating $\sqrt{4.5}$, without a calculator.

In Math 117, you learned how to use differentials to solve problems like this. Observing that $\sqrt{4.5} \approx \sqrt{4} = 2$, we would call $f(x) = \sqrt{x}$, $x_0 = 4$, and $\Delta x = 0.5$, so $\Delta f \approx f'(x_0)\,\Delta x = \frac{1}{2\sqrt{x_0}}\Delta x = \frac{1}{4}(0.5) = \frac{1}{8}$. Therefore, $\sqrt{4.5} \approx 1 + \frac{1}{8} = 2.125$.

This calculation is entirely equivalent to using the linear approximation $P_{1,4}(x)$:

$$\begin{aligned} P_{1,4}(x) &= f(4) + f'(4)(x-4) \\ \implies P_{1,4}(4.5) &= 2 + \tfrac{1}{4}(0.5) \\ &= 2.125 \end{aligned}$$

With Taylor polynomials, though, we have the option of improving on our approximations! A more precise estimate of $\sqrt{4.5}$ can be obtained by calculating $f''(x) = -\frac{1}{4}x^{-3/2}$, $f''(4) = -\frac{1}{32}$, and

$$P_{2,4}(x) = 2 + \frac{1}{4}(x-4) - \frac{1}{64}(x-4)^2.$$

This gives

$$\sqrt{4.5} \approx P_{2,4}(4.5) = 2.125 - \frac{1}{64}(0.5)^2 \approx 2.12109$$

(you can verify with a calculator that this is in fact a closer approximation of the true value).

**Comment:** You might be wondering what the point of this is, since we can, after all, just punch the problem into a calculator. That's a valid objection, but for the time being, take this as an example of the power of Taylor polynomials to make difficult problems easier. I'm

sure you'll agree that calculating $\sqrt{4.5}$ without a calculator is a difficult problem, and yet with Taylor polynomials we can manage it fairly easily!

**Intractable Integrals** How do we evaluate $\int_0^{0.5} \tan^{-1}\left(x^2\right) dx$?

This integral cannot be evaluated exactly, but we can approximate the value. We can show that $P_{6,0}(x)$ for $\tan^{-1}\left(x^2\right)$ is $x^2 - \frac{1}{3}x^6$ (in fact this isn't difficult; we will soon be discussing some shortcuts). This in turn means that

$$\int_0^{0.5} \tan^{-1}\left(x^2\right) dx \approx \int_0^{0.5} \left(x^2 - \frac{1}{3}x^6\right) dx,$$

which is easy to evaluate! This gives an approximate value of 0.041295, which compares very well with the actual value (to 6 d.p.) of 0.041300.

Again, there are calculators and software which can do this for us... but they appeared relatively recently! Again, the point is that with Taylor polynomials we can take a difficult problem, and make it easier.

## 17.2 Maclaurin's Approach

Here's an important question: are the Taylor polynomials unique? In fact they *are* (for each $f$, $x_0$, and $n$). Here's how we know this to be true:

Sometime around 1742, a Scottish mathematician by the name of Colin Maclaurin, who was apparently unaware of Taylor's results, came upon the same formula in a different way (and in fact this is the way Taylor polynomials are derived in most textbooks). Maclaurin's idea was to *start* with a generic $n$th-order polynomial, in powers of $(x - x_0)$, and to then try to determine the appropriate coefficients. So, we let

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n.$$

We want $p(x)$ and its derivatives to have the same values as $f(x)$ and its derivatives at $x_0$ . To start with, if we set $x = x_0$ in $p(x)$, we discover that in order to have $p(x_0) = f(x_0)$, we must have $a_0 = f(x_0)$.

Next, differentiate $p(x)$:

$$p'(x) = a_1 + 2a_2(x - x_0) + 3a_3(x - x_0)^2 + \cdots + na_n(x - x_0)^{n-1}.$$

Setting $x = x_0$ in this expression, and requiring that $p'(x_0) = f'(x_0)$, we discover that we must have $a_1 = f'(x_0)$.

Repeating this, we have

$$p''(x) = 2a_2 + 6a_3(x - x_0) + 12a_4(x - x_0)^2 + \cdots n(n-1)a_n(x - x_0)^{n-2},$$

so if $p''(x_0) = f''(x_0)$, then $2a_2 = f''(x_0)$, i.e. $a_2 = \frac{1}{2}f''(x_0)$.

Continuing in this fashion yields the rest of the coefficients, one by one. We find in general that

$$a_n = \frac{1}{n!}f^{(n)}(x_0),$$

and so $p(x)$ *is* the $n$th-order Taylor polynomial centered at $x_0$!

In recognition of Maclaurin's contribution, we use his name to refer to the simplest special case; a Taylor polynomial centered at zero is referred to as a Maclaurin polynomial. Setting $x_0 = 0$ in Taylor's formula (32), you can see that these have the general form

$$\boxed{P_{n,0}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k}.$$

You can look at the argument above as simply an alternative derivation of the Taylor polynomial formula, but it also guarantees that there is only *one* polynomial of degree $n$ which matches the values of $f$ and its first $n$ derivatives at $x_0$. This is an important realization; it means that if we can find a polynomial which matches the values of $f$ and its derivatives at a point, then that polynomial must be the Taylor polynomial, regardless of how we found it. We can use this to justify some shortcuts. For example, take the linear approximation to $f(x)$ at $x_0 = 0$:

$$f(x) \approx f(0) + f'(0)x \quad \text{for } x \text{ near } 0.$$

If we let $x = t^2$, we discover that

$$f\left(t^2\right) \approx f(0) + f'(0)\, t^2 \quad \text{for } t \text{ near } 0,$$

and this, in fact, the quadratic approximation for the function $f\left(t^2\right)$! How can we be so sure? Well, let's give this function its own name: let $h(t) = f\left(t^2\right)$. Let's also name the approximation: let $Q(t) = f(0) + f'(0)\, t^2$. Differentiation of each gives

$$h'(t) = 2tf'\left(t^2\right), \qquad Q'(t) = 2tf'(0)$$

$$h''(t) = 2f'\left(t^2\right) + 4t^2 f''\left(t^2\right) \qquad Q''(t) = 2f'(0)$$

Evaluating everything at $t = 0$, and assuming that $f''(0)$ exists, we have

$$h(0) = f(0), \text{ and } Q(0) = f(0)$$

$$h'(0) = 0, \text{ and } Q'(0) = 0$$

$$h''(0) = 2f'(0), \text{ and } Q''(0) = 2f'(0)$$

So, the values of $h$, $h'$, and $h''$ are matched by the corresponding values of the polynomial $Q(t)$ at $t = 0$. That means that $Q(t)$ *is* the second-order Maclaurin polynomial of $h(t)$.

*More generally, letting $x = kt^m$ in an $n^{th}$-order Maclaurin polynomial, where $k$ is a real number and $m$ is a positive integer, will always yield the $mn^{th}$-order Maclaurin polynomial of the function $f(kt^m)$, as long as $f^{(mn)}(0)$ exists.*

To see how useful this is, a more specific example may help. Suppose we wish to find an approximation for the function $e^{x^2}$, for values of $x$ near 0. Since we already know that

$$e^t \approx 1 + t + \frac{1}{2}t^2, \qquad \text{for } t \text{ near zero,}$$

we can immediately state that

$$e^{x^2} \approx 1 + x^2 + \frac{1}{2}x^4, \qquad \text{for } x \text{ near zero.}$$

With almost no work at all, we've found the *fourth*-order Maclaurin polynomial for $e^{x^2}$!

**Comment:** You might wonder if we might be able to make other substitutions as well. We can, but the result will not usually be a Taylor polynomial. For example, if we let $t = \sin\theta$ in the quadratic approximation for $e^t$, we find that $e^{\sin\theta} \approx 1 + \sin\theta + \frac{1}{2}\sin^2\theta$ when $\theta$ is small. This gives a reasonable approximation for $e^{\sin\theta}$, but it isn't a *polynomial*!

## 18 The Remainder Theorem for Taylor Polynomials

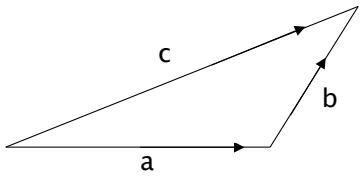
We now turn to the question of accuracy. How good is the approximation $f(x) \approx P_{n,x_0}(x)$? The magnitude of the error is $|f(x) - P_{n,x_0}(x)|$. How large can this be? We won't be able to calculate it exactly, so instead we will try to establish a worst-case scenario; we'll try to find an *upper bound* for it.

To do this, we're going to derive the Taylor polynomial formula in yet another way; we've seen Taylor's derivation of 1715 and Maclaurin's of 1742, and now we'll see what Cauchy did in 1821. The significant difference is that this time we'll obtain the polynomials *and* an expression for the associated error! First, though, we should review one theorem which should have been mentioned in Math 117, but which you have probably not yet had a need to use:

**The Triangle Inequality for Integrals**

- The name "Triangle Inequality" comes from the version for vectors:

$$||\mathbf{a} + \mathbf{b}|| \leq ||\mathbf{a}|| + ||\mathbf{b}||$$




- There is also a Triangle Inequality for scalars:

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

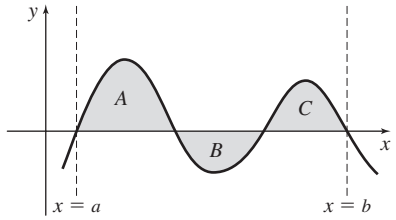
- This can be extended:

$$|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n|$$

- Applying this to Riemann Sums leads to the version we need:

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx \quad \text{(assuming that } a < b\text{)} \tag{33}$$

This has its own simple geometric explanation: if you consider the figure below, in which $A$, $B$, and $C$ are the areas of the regions they label, you'll observe that $\left| \int_a^b f(x)\, dx \right|$ is $A - B + C$, whereas $\int_a^b |f(x)|\, dx$ is $A + B + C$.




---

We're also going to need the Fundamental Theorem of Calculus, Part II. Recall:

$$\boxed{\text{If } F'(x) = f(x)\text{, then } \int_a^b f(x)\, dx = F(b) - F(a)}.$$

This is, in fact, our starting point for our (third) derivation of the Taylor polynomials. We start with a simple change of notation; we can write the above statement as

$$\boxed{\int_{x_0}^{x} f'(t)\, dt = f(x) - f(x_0)}.$$

A slight rearrangement gives

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\, dt,$$

and now you can see why this might be helpful. What we have just written can be interpreted as $P_{0,x_0}(x)$, *plus an expression for the remainder!* Furthermore, that remainder is expressed in terms of an integral, and in Math 117 you learned a couple of methods for rewriting integrals

in different forms... in particular it's integration by parts that will be useful here. If we let $u = f'(t)$ and $dv = dt$, then $du = f''(t)\,dt$ and $v = t$, and we can write

$$\int_{x_0}^{x} f'(t)\,dt = tf'(t)\Big|_{x_0}^{x} - \int_{x_0}^{x} tf''(t)\,dt$$

$$= xf'(x) - x_0 f'(x_0) - \int_{x_0}^{x} tf''(t)\,dt$$

so now $f(x) = f(x_0) + xf'(x) - x_0 f'(x_0) - \int_{x_0}^{x} tf''(t)\,dt$.

Next we need a clever trick: we're trying to get $P_{1,x_0}(x)$, so we need the term $f'(x_0)(x - x_0)$. We *have* $-x_0 f'(x_0)$, but we still *need* $xf'(x_0)$, so we add it (and subtract it... the negative version will become part of the remainder term)! We thus obtain

$$f(x) = f(x_0) + xf'(x) - x_0 f'(x_0) - \int_{x_0}^{x} tf''(t)\,dt + xf'(x_0) - xf'(x_0)$$

$$= f(x_0) + f'(x_0)(x - x_0) + x\left[f'(x) - f'(x_0)\right] - \int_{x_0}^{x} tf''(t)\,dt. \tag{34}$$

This is $P_{1,x_0}(x)$, plus a remainder term, as desired! The remainder term looks a bit unwieldy, but in fact we can tidy it up considerably. Notice that $x\left[f'(x) - f'(x_0)\right]$ can be expressed as $x\int_{x_0}^{x} f''(t)\,dt$ (using the FTC again), and since $x$ is constant with respect to the integral it can be brought inside, giving $\int_{x_0}^{x} xf''(t)\,dt$. This can then be combined with the other integral in (34), and we can now state that

$$f(x) = P_{1,x_0}(x) + \int_{x_0}^{x} (x - t)\,f''(t)\,dt, \tag{35}$$

and we now have an expression for the error involved in using a linear approximation (notice that the error depends upon the *second* derivative of $f$, over the entire interval between the point of tangency, $x_0$, and the point at which we use the approximation, $x$).

To generalize the result to higher orders, we simply need to repeat the integration by parts procedure. In fact, we don't even need any more tricks! In (35), we choose

$$u = f''(t) \qquad dv = (x - t)\,dt$$

$$du = f'''(t)\,dt \quad v = -\frac{1}{2}(x-t)^2$$

and the remainder term becomes

$$\int_{x_0}^{x}(x-t)\,f''(t)\,dt = -\frac{1}{2}(x-t)^2 f''(t)\Big|_{x_0}^{x} + \frac{1}{2}\int_{x_0}^{x}(x-t)^2 f'''(t)\,dt$$

$$= 0 + \frac{1}{2}(x-x_0)^2 f''(x_0) + \frac{1}{2}\int_{x_0}^{x}(x-t)^2 f'''(t)\,dt.$$

Add this to $P_{1,x_0}(x)$, and we have $P_{2,x_0}(x)$, along with a remainder term (involving the *third* derivative of $f$). Repeating this leads to the following:

**The Taylor Theorem with Integral Remainder:**

Suppose that $f$ has $n+1$ derivatives at $x_0$. Then

$$f(x) = \sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x), \tag{36}$$

where

$$R_n(x) = \int_{x_0}^{x}\frac{(x-t)^n}{n!}f^{(n+1)}(t)\,dt.$$

Now, take a deep breath!

We now have an expression for the error in a Taylor polynomial approximation, but we can't evaluate it. As we said at the outset, our goal will be to find an upper bound on its magnitude. The key will be the ability to find an upper bound on the magnitude of the $(n+1)^{\text{st}}$ derivative of $f$ on the interval of interest.

That is, if we can show that $\left|f^{(n+1)}(t)\right| \leq K$ for all values of $t$ between $x_0$ and $x$ (where $K$ is a constant), then we can state that

$$|\text{error}| = |f(x) - P_{n,x_0}(x)|$$

$$= |R_n(x)|$$

$$= \left| \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt \right|$$

Now, if $x_0 < x$, then we can apply the Triangle Inequality for Integrals, giving

$$|R_n(x)| \leq \int_{x_0}^{x} \left| \frac{(x-t)^n}{n!} f^{(n+1)}(t) \right| dt$$

$$= \int_{x_0}^{x} \frac{(x-t)^n}{n!} \left| f^{(n+1)}(t) \right| dt$$

$$\leq \int_{x_0}^{x} \frac{(x-t)^n}{n!} K\, dt = K \int_{x_0}^{x} \frac{(x-t)^n}{n!} dt$$

$$= -K \frac{(x-t)^{n+1}}{(n+1)!} \Bigg|_{x_0}^{x}$$

$$= K \frac{(x-x_0)^{n+1}}{(n+1)!}.$$

If $x_0 > x$, then we need a slightly more careful argument, but we end up with a similar result (we just need an extra negative sign if $n$ is even). Combining the two cases into one, we obtain the result known as...

**Taylor's Inequality:**

The error in using an $n^{\text{th}}$-order Taylor polynomial $P_{n,x_0}(x)$ as an approximation to $f(x)$ satisfies the inequality

$$|R_n(x)| \leq K\frac{|x-x_0|^{n+1}}{(n+1)!},$$

where $\left|f^{(n+1)}(z)\right| \leq K$ for all values of $z$ between $x_0$ and $x$.

(Remember: $x_0$ is the center of the approximation, and $x$ is the location at which we wish to use it.)

**Example:** Suppose we wish to use a $7^{\text{th}}$-order Maclaurin polynomial (that is, a Taylor polynomial centered at zero) to estimate the value of the number $e$.

First, we need to find $P_{7,0}(x)$ for $f(x) = e^x$. This is easy, because all of the derivatives of $e^x$ are just $e^x$, so we have $f(0) = 1$, $f'(0) = 1$, $f''(0) = 1$, ... $f^{(7)}(0) = 1$. Thus we immediately find

$$P_{7,0}(x) = \sum_{k=0}^{7} 1 \cdot \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^7}{7!}.$$

Evaluating $P_{7,0}(1)$ yields the approximation

$$e \approx 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{7!} \approx 2.718253968.$$

Now how accurate might this be?

To use Taylor's Inequality we need to consider $f^{(8)}(x)$, which is of course also $e^x$ (and so $f^{(8)}(z) = e^z$). In this problem $x_0 = 0$ and $x = 1$, so we need to determine how large $e^z$ can be when $z$ is between 0 and 1; clearly on the interval $[0,1]$ we have $|e^z| \leq e$, so we could use this as our value of $K$. However, since our stated goal was to approximate the value of $e$, it's a bit strange to use this same value in our error bound, so let's round up. If we can accept that $e \leq 3$, then we can state that

$$|R_7(x)| \leq 3\frac{|x|^8}{8!}.$$

With $x = 1$ this is $|R_7(x)| \leq \dfrac{3}{8!} = \dfrac{1}{13,440} < 10^{-4}$.

Hence we can justifiably claim that if $e < 3$, then $e = 2.7183$, to 4 decimal places.

**Two Ways to View the Error**

The simplest way to interpret the above result is that on the interval $[0, 1]$, the error is never larger than $1/13440$ in magnitude. Alternatively, if we leave the variable $x$ in the result, then we preserve a bit more information. We have established that

$$|R_7(x)| \leq \frac{x^8}{13,440} \qquad \text{on } [0, 1],$$

and we can see from this that the error shrinks significantly if we use the approximation at a value of $x$ less than 1. This inequality could also be rewritten as

$$P_{7,0}(x) - \frac{x^8}{13,440} \leq e^x \leq P_{7,0}(x) + \frac{x^8}{13,440} \qquad \text{for } x \in [0, 1],$$

from which we can see that we have "pinned" $e^x$ between two polynomials.

**Note:**

Most textbooks give a different form of the Remainder Theorem. You will usually find (36) stated as

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(z)}{(n+1)!}(x - x_0)^{n+1},$$

where $z$ is some number between $x$ and $x_0$, *whose value we cannot determine* (the proof is based on the Mean Value Theorem instead of repeated integration of the FTC).

Don't be too concerned about this difference; both forms of the Remainder Theorem lead to the *same* form of Taylor's Inequality!

**Discussion: Maxima versus Upper Bounds**

If you use the strategies we're going to teach you in the next few weeks, finding a value for $K$ shouldn't be difficult. The function $\left|f^{(n+1)}(z)\right|$ will usually be monotonic on the interval we need to consider, and if it is then we can easily find its maximum (by evaluating it at the endpoints). To use Taylor's Inequality, though we do not actually *need* to find the maximum; an upper bound is all that is required, and if we encounter a function which is not obviously

monotonic, there are a couple of strategies available which can allow us to find an upper bound with little work.

- If we encounter products, we can factor. For example, consider the function $f(x) = \left(x^2+2\right)e^{-x}$, on the interval $[0,2]$. The first factor is increasing on this interval, and we can see that $\left(x^2+2\right) \leq 6$. The second factor is decreasing, and we can see that $e^{-x} \leq 1$. Therefore $f(x) \leq 6$ on $[0,2]$. Of course, you could show that the *maximum* is 2, but 6 is a perfectly valid upper bound, and is easier to find.

- If we encounter sums, we can use the triangle inequality. For example, consider the function $g(x) = x^3 - 2x^2 - 5x + 30$, on the interval $[-3,0]$. It does not look as though $g$ is monotonic, so it is possible that its maximum magnitude occurs somewhere on the interior of the interval. Instead of trying to find it, we could observe that

$$|g(x)| = \left|x^3 - 2x^2 - 5x + 30\right|$$

$$\leq \left|x^3\right| + \left|-2x^2\right| + |-5x| + |30|$$

$$= |x|^3 + 2|x|^2 + 5|x| + 30$$

  Now *this* function is clearly decreasing on $[-3,0]$, so we can set $x = -3$ and conclude that $|g(x)| \leq 27 + 18 + 15 + 30 = 90$. Again, we've overshot the maximum, which happens to be about 32.2, occurring near $x = -0.78$, but finding the maximum requires a lot more work.

In both examples, the upper bound we found was roughly triple the maximum. In an analysis of the accuracy of a Taylor polynomial, this would, at worst, cause us to underestimate the accuracy of our polynomial by one decimal place (and even that would be unlikely).

# 19 Approximation of Integrals using Taylor Polynomials

Suppose we wish to evaluate $\int_0^{0.5} e^{t^2} dt$. It is known that $e^{t^2}$ does not possess a "nice" antiderivative, so there is no way to evaluate this exactly. There are numerical methods available (in fact you might have a calculator which could do it), but what if we wish to examine the behaviour of the *function* $f(x) = \int_0^x e^{t^2} dt$, over an entire interval? For that, we can use Taylor polynomials.

So, let's take this as our challenge: find a polynomial approximation for the function $f(x) = \int_0^x e^{t^2} dt$. Let's say the interval we want is $\left[-\frac{1}{2}, \frac{1}{2}\right]$. We could use our formula for Maclaurin polynomials, and start calculating $f(0)$, $f'(0)$, $f''(0)$, and so on, but it turns out that this is a *really bad idea*! It's a tremendous amount of work, and we have an alternative. The smart way to proceed is to start with the function $g(u) = e^u$, and proceed as follows:

We know (picking $n = 2$ arbitrarily) that, for $u$ close to zero,

$$e^u \approx P_{2,0}(u) = g(0) + g'(0)\,u + \frac{1}{2}g''(0)\,u^2$$

$$= 1 + u + \frac{1}{2}u^2.$$

If we make the substitution $u = t^2$, we obtain the approximation

$$e^{t^2} \approx P_{2,0}\left(t^2\right) = 1 + t^2 + \frac{1}{2}t^4.$$

As discussed at the end of the last section, this must be the 4th-order Maclaurin polynomial for $e^{t^2}$... and we've avoided all the work of having to differentiate $e^{t^2}$ four times! Now, we could call this expression $P_{4,0}(t)$ (for the function $e^{t^2}$), but for clarity let us keep the original name for the function; it's $P_{2,0}(u)$ for $e^u$, *evaluated at* $t^2$.

This in turn means that

$$\int_0^x e^{t^2} dt \approx \int_0^x P_{2,0}\left(t^2\right) dt = \int_0^x \left(1 + t^2 + \frac{1}{2}t^4\right) dt$$

$$= x + \frac{x^3}{3} + \frac{x^5}{10}.$$

For the more specific problem where $x = 0.5$, this gives

$$\int_0^{0.5} e^{t^2} dt \approx 0.54479166\overline{6}.$$

We hope you'll agree that finding the approximation is not all that hard. Through substitutions, integration, and differentiation, we can manipulate the polynomial approximations of simple functions to obtain approximations for more complicated ones. As you'll see, if we're looking for a Maclaurin polynomial, it's usually possible to start with just one of five functions: $e^x$, $\sin x$, $\cos x$, $\frac{1}{1-x}$, and the function $(1+x)^k$, where $k$ can be any constant.

Now, how accurate is our approximation? To answer this question, we'll go back to the beginning again.

We know that

$$e^u = P_{2,0}(u) + R_2(u),$$

where

$$|R_2(u)| \leq K\frac{|u|^3}{3!},$$

and

$$\left|f^{(3)}(z)\right| \leq K \quad \text{for } z \text{ between 0 and } u.$$

Since $f(u) = e^u$, we have $\left|f^{(3)}(z)\right| = e^z$, but we need to know what interval to consider. We need to consider the string of variables we're going to go through to obtain the approximation:

- $z$ must lie between 0 and $u$.
- we are going to make the substitution $u = t^2$.
- $t$ is our variable of integration, and must take on all the values between 0 and $x$.
- we originally stated that we wanted to consider $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$.

Working backwards, then,

$$\begin{aligned} x \in \left[-\frac{1}{2}, \frac{1}{2}\right] &\implies t \in \left[-\frac{1}{2}, \frac{1}{2}\right] \\ &\implies u \in \left[0, \frac{1}{4}\right] \\ &\implies z \in \left[0, \frac{1}{4}\right]. \end{aligned}$$

On this interval, we can state that $\left|f^{(3)}(z)\right| = e^z \leq e^{1/4} < 2$ (rounding up for convenience). This can be used as our value for $K$, so we can restate our initial approximation as

$$e^u = P_{2,0}(u) + R_2(u), \quad \text{where } |R_2(u)| \leq 2\frac{|u|^3}{3!} \qquad \text{for } u \in \left[0, \frac{1}{4}\right].$$

Next, we replace $u$ with $t^2$, to obtain

$$e^{t^2} = P_{2,0}\left(t^2\right) + R_2\left(t^2\right), \quad \text{where } \left|R_2\left(t^2\right)\right| \leq \frac{t^6}{3} \qquad \text{for } t \in \left[-\frac{1}{2}, \frac{1}{2}\right].$$

Finally, we integrate from 0 to $x$:

$$\int_0^x e^{t^2} dt = \int_0^x P_{2,0}\left(t^2\right) dt + \int_0^x R_2\left(t^2\right) dt,$$

where

$$\begin{aligned}\left|\int_0^x R_2\left(t^2\right) dt\right| &\leq \int_0^x \left|R_2\left(t^2\right)\right| dt \\ &\leq \int_0^x \frac{1}{3} t^6 dt \\ &\leq \frac{1}{21}x^7, \qquad \text{for } x \in \left[-\frac{1}{2}, \frac{1}{2}\right].\end{aligned}$$

A technical point here: the Triangle Inequality for integrals requires that the upper limit of integration be greater than the lower limit, so the last step of this calculation is only valid for $x > 0$. If $x < 0$, we must first rewrite $\int_0^x$ as $-\int_x^0$. Proceeding this way gives $\left|\int_0^x R_2\left(t^2\right) dt\right| \leq -\frac{1}{21}x^7$ (try verifying this), so as a general result we can say that

$$\left|\int_0^x R_2\left(t^2\right) dt\right| \leq \frac{1}{21}|x|^7, \qquad \text{for } x \in \left[-\frac{1}{2}, \frac{1}{2}\right].$$

Summarizing, we have established that

$$\int_0^x e^{t^2} dt = x + \frac{x^3}{3} + \frac{x^5}{10} \pm \frac{x^7}{21}, \qquad \text{for } x \in \left[-\frac{1}{2}, \frac{1}{2}\right],$$

and specifically,

$$\int_0^{0.5} e^{t^2} dt = 0.054479 \pm 0.0004$$

(so, to three decimal places, the value is 0.545).

One final comment about notation: the symbol "$\pm$" is ambiguous. In mathematics it is usually used to denote one value *or* the other (for example, in the quadratic formula: $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ represents exactly two values). In the current context we're using it to include a continuous range of values between two extremes. This is common in the sciences, but some mathematicians would frown upon it!

## 20 Infinite Series

So far we've been suggesting that the accuracy of our Taylor polynomial approximations should improve when we incorporate more terms. This certainly appears to be true in the examples we've discussed so far, and in fact for some functions it holds true even when we consider values of $x$ which are at great distances from the center of the polynomial. As an example, consider $f(x) = \sin x$:

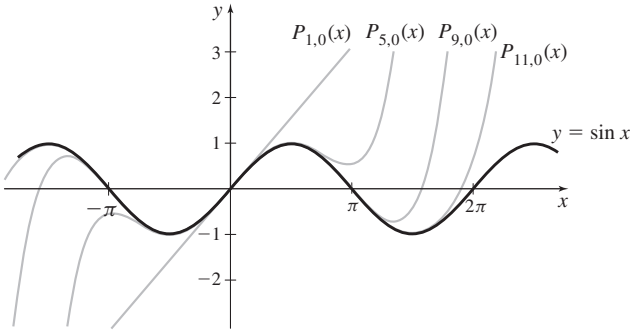
$$\begin{aligned}
f(x) &= \sin x & f(0) &= 0 \\
f'(x) &= \cos x & f'(0) &= 1 \\
f''(x) &= -\sin x & f''(0) &= 0 \\
f'''(x) &= -\cos x & f'''(0) &= -1 \\
f^{(4)}(x) &= \sin x & f^{(4)}(0) &= 0 \\
&\vdots & &\vdots \\
& & f^{(2n+1)}(0) &= (-1)^n,\ f^{(2n)}(0) = 0
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow \qquad P_{2n+1,0}(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\
&= \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!}.
\end{aligned}$$

Even if we consider a relatively large value of $x$, say $x = 3$, we find that with enough terms we still obtain a reasonable approximation of the value of the function:

$$P_{1,0}(3) = 3$$
$$P_{3,0}(3) = -1.5$$
$$P_{5,0}(3) = 0.525$$
$$P_{7,0}(0) = 0.091$$
$$P_{9,0}(0) = 0.145$$
$$P_{11,0}(0) = 0.141$$
$$P_{13,0}(0) = 0.141$$
$$\vdots$$

In fact, it turns out that we'll obtain a sequence converging to the correct value for *any* choice of $x$. The figure below shows how the Taylor polynomial approximations match the curve $y = \sin x$ over longer and longer intervals as $n$ increases.




However, things don't *always* happen this way! Consider the function $f(x) = \dfrac{1}{1+x}$. This time we find

$$\begin{aligned} f(x) &= \tfrac{1}{1+x} & f(0) &= 1 \\ f'(x) &= -\tfrac{1}{(1+x)^2} & f'(0) &= -1 \\ f''(x) &= \tfrac{2}{(1+x)^3} & f''(0) &= 2 \\ f'''(x) &= -\tfrac{6}{(1+x)^4} & f'''(0) &= -6 \\ &\vdots & &\vdots \\ f^{(n)}(x) &= (-1)^n \tfrac{n!}{(1+x)^{n+1}} & f^{(n)}(0) &= (-1)^n n! \end{aligned}$$

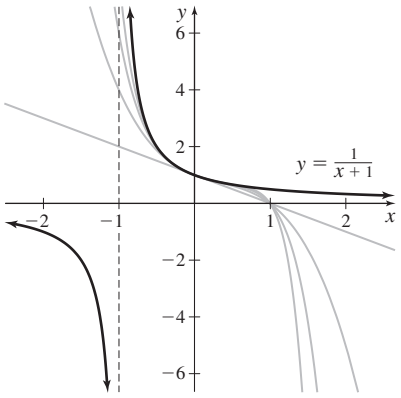
$$\Longrightarrow \qquad P_{n,0}(x) = 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n$$

$$= \sum_{k=0}^{n} (-1)^k x^k.$$

These approximations work quite well if $x$ is small, but if we take, say, $x = 2$ (where the output is $1/3$) we find this:

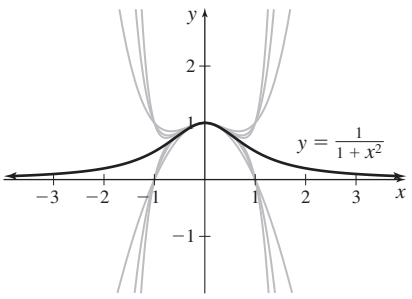
$$P_{0,0}(2) = 1$$
$$P_{1,0}(2) = -1$$
$$P_{2,0}(2) = 3$$
$$P_{3,0}(2) = -5$$
$$\vdots$$

These approximations continue to get *worse* as we incorporate more terms! Here are the graphs of the first few odd polynomials:




So, what's gone wrong? Well, we do know that we should expect to encounter some problems with this function, since it has a discontinuity at $x = -1$ (so obviously the polynomial approximation will break down near that point). Why, though, should this affect the "approximations" at $x = 2$?

In fact the situation is even more complicated; the Taylor polynomials for $\dfrac{1}{1+x^2}$ and $\tan^{-1}(x)$ exhibit similar behaviour, and these functions don't have any discontinuities.[23]




[23] At least, not if we restrict our view to real numbers!

### Further Discussion

At the moment the only tool we have for analyzing the error in Taylor polynomial approximations is Taylor's Inequality, so let's see if it can shed any light on the problem. Recall that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_n(x),$$

where $|R_n(x)| \leq K\dfrac{|x - x_0|^{n+1}}{(n+1)!}$, with $\left|f^{(n+1)}(z)\right| \leq K$ for all values of $z$ between $x$ and $x_0$.

In the case of $\sin x$, with $x_0 = 0$, we have

$$\sin x = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + R_{2n+1}(x),$$

with $|R_{2n+1}(x)| \leq K\dfrac{|x|^{2n+2}}{(2n+2)!}$, where $K$ is an upper bound on the magnitude of the $(2n+2)^{\text{nd}}$ derivative of $\sin x$. Of course, for this function we can use $K = 1$, on *any* interval, for *any* value of $n$. This gives

$$|R_{2n+1}(x)| \leq \frac{x^{2n+2}}{(2n+2)!}.$$

Now, what happens as $n \to \infty$? Well, the sequence $\dfrac{x^{2n+2}}{(2n+2)!}$ is related to the sequence $\dfrac{x^n}{n!}$ (we call it a *subsequence*), and it can be shown that $\lim\limits_{n\to\infty} \dfrac{x^n}{n!} = 0$ (we'll see an easy argument for this later on). It follows that $\dfrac{x^{2n+2}}{(2n+2)!} \to 0$ as $n \to \infty$, and hence by the Squeeze Theorem we have also that $R_{2n+1}(x) \to 0$ as $n \to \infty$, *for* ***any*** *x!* This tells us that for the sine function, we can always improve our approximations by adding more terms.

### Taylor Series

At this point we're ready to formalize an idea which may have already occurred to you: in the limit as $n \to \infty$, our approximations may become exact! We have established that

$$\sin x = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + R_{2n+1}(x)$$

and $\lim_{n\to\infty} R_{2n+1}(x) = 0$. Clearly it is also true that $\lim_{n\to\infty} \sin x = \sin x$ (since this is independent of $n$), and so the limit of the series must also exist (because the series is the difference of two convergent sequences).

Therefore, upon taking limits we discover that

$$\sin x = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!}, \qquad \text{for all } x. \tag{37}$$

We usually write this in a shorthand form (using the same convention that we use for improper integrals):

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}. \tag{38}$$

This is called the *Taylor Series centered at zero* of $\sin x$ (or simply the *Maclaurin Series* of $\sin x$).

**WARNING: This is *not* exactly a sum; it's a limit of sums. The notation in Equation 38 *is defined to mean* what is written in Equation 37. As you will see, there are cases where the usual rules for sums will not apply!**

To make this more general, we can state that since

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_n(x),$$

we must have

$$f(x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \lim_{n\to\infty} R_n(x).$$

Therefore IF $\lim_{n\to\infty} R_n(x_0) = 0$, then $f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$. Otherwise, we may still

be able to calculate the Taylor series, but it won't be equal to $f(x)$ (and so it probably won't be of much use).

So, we've seen why the approximations for $f(x) = \sin x$ work so well, but what happens with $g(x) = \frac{1}{1+x}$? In this case, Taylor's Inequality tells us that

$$\frac{1}{1+x} = \sum_{k=0}^{n} (-1)^k x^k + R_n(x),$$

where

$$|R_n(x)| \leq K \frac{|x|^{n+1}}{(n+1)!},$$

with

$$\left| f^{(n+1)}(z) \right| \leq K \qquad \text{for all } z \text{ between 0 and } x.$$

This time, though, $f^{(n+1)}(z) = (-1)^n \frac{(n+1)!}{(1+z)^{n+2}}$, and it is *not* as easy to find an upper bound as it was for $\sin x$. In fact, you can see that if we consider the interval $(-1, 1)$, $f^{(n+1)}(z)$ is not bounded at all, so we cannot use a single value of $K$ for this entire interval (we would need different values of $K$ for different subintervals). Even after we restrict ourselves to an interval which avoids the discontinuity at -1 we find that $K$ must depend on $n$. For example, if we use the interval $(0, a)$ for some constant $a > 0$, we can use $K = (n+1)!$. There's nothing wrong with this, but the conclusion is that $|R_n(x)| \leq |x|^{n+1}$. Since this will approach zero only if $|x| < 1$ this tells us that $\frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k$ for $x \in (0, 1)$. However, we need more work to figure out what happens outside of this interval.

In short, Taylor's Inequality is answering some of our questions, but it is not giving us the answers easily! In fact, for most functions other than $\sin x$ and $\cos x$ it will be quite difficult to analyze $R_n(x)$, so we will instead turn to ways of analyzing properties of the infinite series themselves. That is, rather than showing that $R_n(x) \to 0$ as $n \to \infty$, we'll discuss some tools for determining directly whether $P_{n,x_0}(x) \to f(x)$, by investigating the behaviour of the infinite series of *constants* generated when we assign values to $x$.

## 21 Convergence of Infinite Series

In the previous section we made an attempt to discover when it might be true that $f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$ (that is, when a function might be equal to its Taylor series). We were able to use Taylor's Inequality to show that it is true for $\sin x$ (and similarly, we could show it to be true for $\cos x$ as well), but encountered difficulty in using the same approach for other functions. So, we will try a different approach: we'll analyze the behaviour of the series itself.

Consider what happens when we assign a value to $x$; we obtain an infinite series of *constants*! For example, we've established that $\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$. Setting $x = 1$, we obtain the expression

$$\sin 1 \;=\; 1 - \frac{1}{6} + \frac{1}{120} - \frac{1}{5040} \pm \ldots.$$

Does this really makes sense, though? *Is it really possible to add up infinitely many numbers and obtain a finite result*?

In some cases this is obviously nonsensical. For example, if we write

$$1 + 1 + 1 + 1 + 1 + \ldots$$

you will probably agree that this cannot have a sum - or you might prefer to say that the sum is infinite.

On the other hand, there is one usage of infinite series that is quite familiar; you might just not have given it much thought. Consider the expression

$$1 + 0.1 + 0.01 + 0.001 + 0.0001 + \ldots.$$

This is precisely what we mean when we write the number $1.111\ldots$, which you might recognize as being $10/9$. So, we've already determined that some infinite series represent finite numbers (they *converge*), while others do not. That leads us to the key question: *can we distinguish between the two*?

If we can answer this question, we'll know when our Taylor polynomial approximations "work". However, some caution is required; it's easy to make mistakes when working with infinite series, as the following calculation shows:

$$
\begin{array}{rcccccccccc}
0 &=& 0 &+& 0 &+& 0 &+& 0 &+& \ldots \\
&=& (1-1) &+& (1-1) &+& (1-1) &+& (1-1) &+& \ldots \\
&=& 1-1 &+& 1-1 &+& 1-1 &+& 1-1 &+& \ldots \\
&=& 1 &+& (-1+1) &+& (-1+1) &+& (-1+1) &+& \ldots \\
&=& 1 &+& 0 &+& 0 &+& 0 &+& \ldots \\
&=& 1 & & & & & & & &
\end{array}
$$

Legend has it that, around 1703, in letters to contemporary mathematicians, an Italian monk by the name of Guido Grandi (often called Guido Ubaldus) presented this as proof of the existence of God, since it suggested that it is possible to produce something out of nothing! Whether or not this is what he really meant by what he wrote isn't clear, but we do know that the brightest minds of the day were unsure how to explain what the problem was. Gottfried Leibniz at least recognized that the problem was in the third line above; the sums above that are clearly zero, and the sums below it are clearly 1. Leibniz suggested that the sum $1 - 1 + 1 - 1 + 1 - 1 + \ldots$ should be assigned the value $1/2$, since it appeared to have equal probability of giving a value of 0 or 1!

In today's terminology we say that the first two sums *converge* to zero, the last three *converge* to 1, while the one in between simply *does not converge* (we say it *diverges*). We also have a precise definition of what we mean by this.

**Definition:** An *infinite series* (or just *series*) of constants $a_k$ is defined as a limit of finite series:

$$\sum_{k=0}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=0}^{n} a_k.$$

In other words, given a sequence of numbers $\{a_k\}$, we construct the *sequence of partial sums*, $\{s_n\}$:

$$
\begin{aligned}
s_0 &= a_0 \\
s_1 &= a_0 + a_1 \\
s_2 &= a_0 + a_1 + a_2 \\
\vdots &= \quad \vdots \\
s_n &= a_0 + a_1 + a_2 + \cdots + a_n.
\end{aligned}
$$

If the sequence $\{s_n\}$ converges (to "$s$", say), that is if $\lim_{n\to\infty} s_n = s$, then we say that the series $\sum_{k=0}^{\infty} a_k$ *converges*, and that its sum is $s$. Otherwise, we say it *diverges*.

A few comments might be helpful here:

1. Our definition is just a precise mathematical statement of a very intuitive idea. Consider the series $\sum_{k=0}^{\infty} \frac{1}{2^k}$, which is $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots$. If you take your calculator and start adding, you'll see the numbers

$$\begin{array}{rcl} 1 & & \\ +1/2 & = & 1.5 \\ +1/4 & = & 1.75 \\ +1/8 & = & 1.875 \\ +1/16 & = & 1.9375 \\ +1/32 & = & 1.96875 \\ +1/64 & = & 1.984375 \\ \vdots & & \vdots \end{array}$$

   and if you keep going you'll find that the sequence of numbers on the right gets closer and closer to 2, so it "feels" as though that should be the sum. Well, that list of numbers on the right is exactly what we've used to define the sum of a series; it's the sequence of partial sums!

2. In everyday speech, the words "sequence" and "series" are often interchangeable, but for mathematical usage we've introduced a clear and important distinction: a *sequence* is simply a list of numbers, while a *series* is what we obtain when we add those numbers together.

3. Notice that every series is associated with two different sequences: the sequence of components $\{a_k\}$ and the sequence of partial sums $\{s_n\}$.

4. For simplicity, we've written the definition with the index $k$ starting at zero (and most textbooks do the same). However, the index could start at any integer, so to be more general we could have defined $\sum_{k=q}^{\infty} a_k$, for $q \in \mathbb{Z}$. This isn't actually necessary, though,

since we can always "re-index" a series if we wish:

$$\sum_{k=q}^{\infty} a_k = \sum_{k=0}^{\infty} a_{(k+q)}$$

(you should be able to see that both of these expressions represent the series $a_q + a_{(q+1)} + a_{(q+2)} + \ldots$). Also, if the question we're trying to answer is *whether a series converges or not*, then it really doesn't matter where the index begins! Adding a few numbers to (or removing some from) the beginning of a series will only affect the *value* of the sum, not whether it exists (compare $1 + 0.1 + 0.01 + \ldots$ to $0.1 + 0.01 + 0.001 + \ldots$ for example). For this reason we'll often neglect the initial index value entirely when we're stating general results, and just write $\sum a_k$.

Now that we have a precise definition, we can avoid making mistakes of the kind that Guido Grandi made. If you consider again the series

$$1 - 1 + 1 - 1 + 1 - 1 + \ldots$$

you'll observe that the sequence of partial sums is

$$1, 0, 1, 0, 1, 0, \ldots$$

which has no limit. Therefore the series is divergent, and cannot be said to be equal to 0 or 1, or anything else!

## The Convergence Tests

Ideally, we'd like to be able to look at a series and recognize immediately whether it converges or not, but it isn't quite as straightforward as might be hoped. There are no less than eight theorems to introduce! The good news is that in practice, one of them stands out as being *almost* entirely sufficient by itself, and once you've had some practice you'll be able to use many of the others successfully with very little work. Seven of these theorems are referred to as the "convergence tests". We'll start with the one which is not.

## 1. Geometric Series

A geometric series has the form

$$\sum_{k=0}^{\infty} ar^k = a + ar + ar^2 + ar^3 + \dots.$$

You may have seen finite geometric series in high school, and you might recall that for these we have a formula for the sum. This allows us to state that

$$\sum_{k=0}^{\infty} ar^k = \lim_{n\to\infty} \sum_{k=0}^{n} ar^k$$

$$= \lim_{n\to\infty} \frac{a\left(1 - r^{n+1}\right)}{1-r}.$$

So, what happens as $n \to \infty$?

- If $|r| < 1$, the sequence $r^n$ approaches zero, and so we can conclude that $\displaystyle\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$.
- If $|r| > 1$, the sequence $r^n$ diverges, and so the series diverges as well.
- If $r = 1$ the sequence $r^n = 1^n$ converges to 1, and our formula gives us a sum of $\frac{0}{0}$. That's not much use, but looking at the series directly we have $\sum 1 = 1+1+1+1+\dots$, which obviously diverges.
- If $r = -1$ the sequence $r^n = (-1)^n$ diverges, and so the series does as well. In fact, this is Guido Grandi's infamous series again!

Summarizing, we see that

$$\boxed{\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r} \qquad \text{if } |r| < 1, \text{ and is divergent otherwise.}}$$

**Examples**

- $\displaystyle\sum_{k=0}^{\infty} 10^k$ diverges, since $10 > 1$. (This series is $1 + 10 + 100 + 1000 + \dots$, and its sequence of partial sums is $\{1, 11, 111, 1111, \dots\}$).
- $\displaystyle\sum_{k=0}^{\infty} \left(\frac{1}{10}\right)^k$ converges, since $\frac{1}{10} < 1$. Furthermore, its sum is $\dfrac{1}{1-\frac{1}{10}} = \dfrac{1}{9/10} = 10/9$. We discussed this one briefly earlier; using the definition of convergence we saw that

the sequence of partial sums is $\{1, 1.1, 1.11, 1.111, 1.1111, \ldots\}$, which converges to the number $1.\overline{111}$ (which is the decimal representation of $10/9$).

- $\sum_{k=0}^{\infty} \left(\frac{4}{5}\right)^k$ converges, to the sum $\frac{1}{1-\frac{4}{5}} = 5$. With this example, we hope you'll see the usefulness of the theorem; if you were to simply start writing out the sequence of partial sums, you'd find $\{1, 1.8, 2.44, 2.952, 3.3616, \ldots\}$, and you'd probably need a fair amount of time on your hands to satisfy yourself that the limit exists!

**Note:** Geometric series won't always appear in precisely the form $\sum_{k=0}^{\infty} ar^k$, and so you may find it helpful to think of the formula for the sum not as $\frac{a}{1-r}$, but as $\frac{\text{"first term"}}{1-\text{"common ratio"}}$. For example, the series $\sum_{k=3}^{\infty} 4\left(\frac{1}{3}\right)^{k-1}$ is geometric, since its terms are defined by the common ratio $\frac{1}{3}$ (it's $\left\{\frac{4}{9} + \frac{4}{27} + \frac{4}{81} + \ldots\right\}$). To use the formula $\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$ we would have to re-index the series as $\sum_{k=0}^{\infty} \frac{4}{9}\left(\frac{1}{3}\right)^k$. There are situations in which re-indexing will be a useful skill, but we can avoid it here; the first term in our series is $\frac{4}{9}$, and the common ratio is $\frac{1}{3}$, so the sum is

$$\frac{\frac{4}{9}}{1-\frac{1}{3}} = \frac{4/9}{2/3} = \frac{2}{3}.$$

Actually, this is not the first time we've encountered a geometric series! Recall that we found the Maclaurin series of $f(x) = \frac{1}{1+x}$ to be $\sum_{n=0}^{\infty} (-1)^n x^n$. Replacing $x$ with $-x$, we find that $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$, and this is precisely the formula we've just discussed (just set $a = 1$ and $r = x$). We now know that this Maclaurin series is valid for values of $x$ between $-1$ and 1.

Unfortunately, if we are given a series which is not geometric, it is unlikely that we will be able to determine its sum exactly[24]. However, the following theorems should at least allow us to determine *if* a sum exists (that is, to determine whether the series is convergent or not). There are also some ways in which we can approximate sums, to any desired precision, but we won't discuss all of these.

[24] There are a handful of non-geometric patterns we might recognize. For example, you might be able to see that $\sum_{n=0}^{\infty} \frac{(-1)^n 2^{2k+1}}{(2k+1)!}$ is just $\sin(2)$. What about $\sum_{n=1}^{\infty} \frac{1}{2^n n!}$?

## 2. The Test for Divergence (a.k.a. the Nth Term Test)

$$\boxed{\text{If } \lim_{k\to\infty} a_k \neq 0, \text{ then } \sum a_k \text{ diverges.}}$$

**Examples:**

- Consider the series $\displaystyle\sum_{k=1}^{\infty} \frac{k^2+2}{4k^2-k}$. We can tell immediately that this is divergent, because $\dfrac{k^2+2}{4k^2-k} \to \dfrac{1}{4}$ as $k \to \infty$. The logic behind the theorem should be clear; eventually this particular series behaves more or less like the series $\dfrac{1}{4}+\dfrac{1}{4}+\dfrac{1}{4}+\cdots$, so the partial sums continue to grow, infinitely.
- Similarly, the series $\displaystyle\sum_{k=1}^{\infty} e^{1/k}$ must diverge, since $e^{1/k} \to 1$ as $k \to \infty$.

There is an absolutely critical point to be made here: the theorem is not an if-and-only-if statement; it says nothing at all about what might happen if the limit IS zero!

That is, if $\lim_{k\to\infty} a_k = 0$, we may have either convergence or divergence. To demonstrate this, consider the two series $\sum \dfrac{1}{k^2}$ and $\sum \dfrac{1}{k}$. In both cases, the sequence of individual terms approaches zero as $k$ increases. The first series happens to converge (as we'll explain shortly), but the second one does not! This is such an important discovery that the series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$ gets a special name; it is known as the *harmonic series*. The terms in the sequence $\left\{\dfrac{1}{k}\right\}$ approach zero, and yet the harmonic series diverges!

There is a clever little proof of this result; we take the $2^k$th partial sum, and group the terms in a certain way (we double the number of terms in each successive group):

$$\begin{aligned} s_{2^k} &= 1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)+\left(\frac{1}{9}+\ldots+\frac{1}{16}\right)+\ldots+\left(\ldots+\frac{1}{2^k}\right) \\ &> 1+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)+\left(\frac{1}{16}+\ldots+\frac{1}{16}\right)\ldots+\left(\frac{1}{2^k}\ldots+\frac{1}{2^k}\right) \\ &= 1+\frac{1}{2} \quad +\frac{1}{2} \quad +\frac{1}{2} \quad +\frac{1}{2} \quad \ldots+\frac{1}{2} \\ &= 1+\frac{k}{2} \end{aligned}$$

Letting $k$ increase, we see that the partial sums can be made as large as we wish. That is, the series diverges!

You should probably take some time to think about this. It is tempting to think that because the individual terms approach zero, they should eventually be negligible. However, there are infinitely many of them, and if they don't shrink quickly enough, they can still accumulate without bound!

Perhaps it will help if we state the Test for Divergence in another way:

For $\sum a_k$ to converge, the condition $a_k \to 0$ as $k \to \infty$ is *necessary*, but not *sufficient*!

### 3. The Integral Test (applicable only to series with $a_k > 0$ )

Consider a series $\sum_{k=k_0}^{\infty} a_k$. Let $f$ be a function which is continuous, positive, and decreasing on $(k_0, \infty)$, with $f(k) = a_k$ for all $k \geq k_0$.

$\sum_{k=k_0}^{\infty} a_k$ converges if and only if $\int_{k_0}^{\infty} f(x)\,dx$ converges.

In most cases, the appropriate function $f$ can be obtained simply by replacing $k$ with $x$. The logic behind the theorem is easy to illustrate; consider the two figures below. We use the values of $f$ at the integers to define the heights of rectangles of base width 1. The area of each rectangle is therefore $f(k)$, and the sum of the series (if it exists) corresponds to the total shaded area. By positioning the rectangles to the left of their defining integers, as in the first figure, we can see that if the area below the curve $y = f(x)$ is finite, then the total area of the rectangles must also be finite (that is, the series must converge). By positioning them to the right instead, we can see that the converse must also be true: if the series converges then the integral must converge as well.

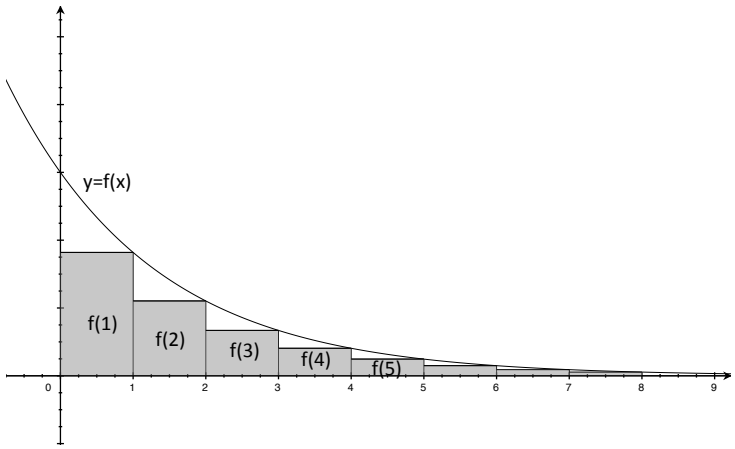



Figure 41:

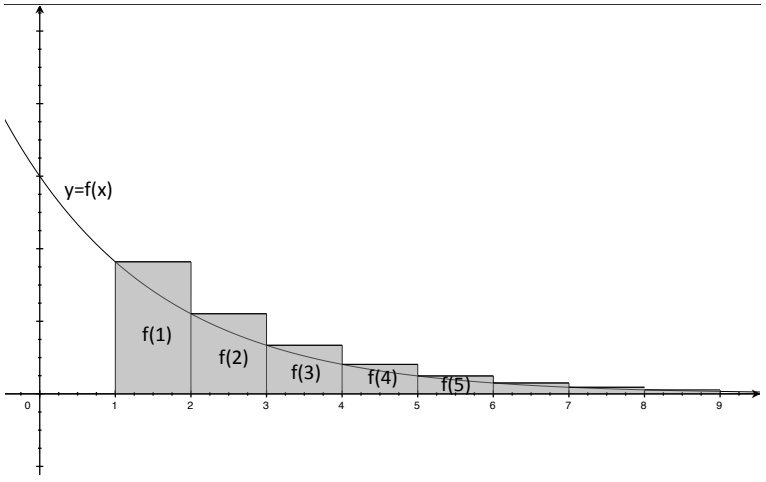



Figure 42:

**Examples**

- Consider the series $\displaystyle\sum_{k=2}^{\infty} \frac{1}{k\,(\ln k)^2}$. To apply the integral test, we consider the improper integral $\displaystyle\int_2^{\infty} \frac{dx}{x\,(\ln x)^2}$. Evaluating this (and letting $u = \ln x$ to do so), we find

$$\lim_{t\to\infty} \int_2^{\infty} \frac{dx}{x\,(\ln x)^2} = \lim_{t\to\infty} \int_{\ln 2}^{\ln t} \frac{du}{u^2}$$

$$= \lim_{t\to\infty} \; -\frac{1}{u}\Big|_{\ln 2}^{\ln t}$$

$$= \lim_{t\to\infty} \left(\frac{1}{\ln 2} - \frac{1}{\ln t}\right)$$

$$= \frac{1}{\ln 2}.$$

  Since the integral converges, we may conclude that the series converges as well. Note, however, that this does NOT mean that the sum of the series is $\frac{1}{\ln 2}$. The value will generally be different (although it is possible to use similar integrals to *approximate* the sum of the series).

- Now consider the series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$. If we consider the corresponding integral, we have

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

$$= \lim_{t\to\infty} 2\sqrt{x}\Big|_1^{t}$$

$$= \lim_{t\to\infty} \left(2\sqrt{t} - 2\right) = \infty.$$

The integral diverges, and therefore the series does as well.

This second example is actually a special case of a family of series known as "p-series". By using the Test for Divergence or the Integral Test, we can analyze any series of the form $\sum \frac{1}{k^p}$, and the result can be considered as a test in its own right:

## 4. P-Series

$$\boxed{\sum \frac{1}{k^p} \text{ converges if } p > 1, \text{ and diverges if } p \leq 1.}$$

This is an extremely useful result. Like the theorem for Geometric series, it requires no work; we can distinguish between convergent and divergent p-series on sight.

**Examples:**

- $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}, \quad \sum_{k=1}^{\infty} \frac{1}{k^3}, \quad \sum_{k=1}^{\infty} \frac{1}{k^{1.5}},$ etc. all converge.
- $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}, \quad \sum_{k=1}^{\infty} \frac{1}{\sqrt[3]{k}},$ etc. all diverge.

## 5. The Comparison Test (applicable only to series with $a_k > 0$ )

Suppose we are given a series $\sum a_k$ (with all terms positive). If we can identify a second series $\sum b_k$ such that $a_k \leq b_k$ for all $k$, and $\sum b_k$ converges, then $\sum a_k$ also converges. Similarly, if $a_k \geq b_k$ and $\sum b_k$ diverges, then $\sum a_k$ also diverges.

The series we choose for comparison, $\sum b_k$, needs to be a series whose behaviour we already know. Normally, it will be a geometric series or a p-series.

**Examples:**

- Consider $\sum_{k=3}^{\infty} \frac{\ln k}{k}$. We know that $\sum_{k=3}^{\infty} \frac{1}{k}$ diverges (this is the harmonic series), and we also know that $\frac{\ln k}{k} > \frac{1}{k}$ (at least when $k \geq 3$). Therefore $\sum_{k=3}^{\infty} \frac{\ln k}{k}$ also diverges.

- What about $\sum_{k=2}^{\infty} \frac{\ln k}{k}$? Well, this is just the previous series with the term $\frac{\ln 2}{2}$ added, so it must diverge as well. We can ignore any number of initial terms if necessary; in fact in each case where a test says "for all $k$" we could replace that phrase with "for all $k$ larger than some number".

- Consider $\sum_{k=1}^{\infty} \frac{1}{k^2+2}$. We know that $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges (it's a p-series), and we know that $\frac{1}{k^2+2} < \frac{1}{k^2}$, for all $k$. Therefore the series $\sum_{k=1}^{\infty} \frac{1}{k^2+2}$ converges.[25]

## 6. The Limit Comparison Test (the "LCT", applicable only to series with $a_k > 0$ )

To motivate the need for this next test, consider the series $\sum_{k=2}^{\infty} \frac{1}{k^2-2}$. When $k$ becomes large this series should also behave like $\sum_{k=1}^{\infty} \frac{1}{k^2}$, but the inequality needed for the Comparison Test is not satisfied $\left(\frac{1}{k^2-2} \not< \frac{1}{k^2}\right)$. It might still "feel" to you as though the series should be convergent, though, and in fact we do have a theorem which can be used to prove this:

| If $\lim_{k\to\infty} \frac{a_k}{b_k} = L$, where $L$ is a nonzero constant $(0 < L < \infty)$, then $\sum a_k$ and $\sum b_k$ either both converge or both diverge. |
|---|

**Examples:**

- The series $\sum_{k=2}^{\infty} \frac{1}{k^2-2}$ converges, because $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges, and

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

[25] We do not know what the sum is, but we can at least say that it must be less than $\pi^2/6$. Can you see why?

- The series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{2+3\sqrt{k}}$ diverges, because $\displaystyle\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ diverges, and

$$\lim_{k\to\infty} \frac{\frac{1}{2+3\sqrt{k}}}{\frac{1}{\sqrt{k}}} = \lim_{k\to\infty} \frac{\sqrt{k}}{2+3\sqrt{k}} = \frac{1}{3}.$$

  The idea here is to pick the geometric series or p-series which most resembles the series in question when $k$ is large.

**Comment:** It should be possible to get $L = 1$ if we pick the right series (in our second example above, we could have used $\sum \frac{1}{3\sqrt{k}}$ instead of $\sum \frac{1}{\sqrt{k}}$ for our comparison). However, this isn't necessary; the reason the theorem uses "$L$" instead of "1" is to save us having to think about this; any multiple of the "right" series will do.

## 7. The Alternating Series Test (the "AST", a.k.a. the Leibniz Test)

Consider the series $a_0 - a_1 + a_2 - a_3 + \ldots \quad = \displaystyle\sum_{k=0}^{\infty} (-1)^k a_k$, where $a_k > 0$ for every $k$.
If $\lim\limits_{k\to\infty} a_k = 0$ *and* the sequence $\{a_k\}$ is decreasing[a], then the series converges.

[a] As we discussed for the Comparison Test, all we really need is that the sequence *eventually* be decreasing (i.e. $a_{k+1} < a_k$ for all $k$ greater than some integer).

**Examples:**

- Consider $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}}$. First of all, you should realize that this is not exactly a p-series, because it contains negative terms. It is, however, alternating, and since the sequence $\left\{\frac{1}{\sqrt{k}}\right\}$ is clearly decreasing (for all $k$), and $\lim\limits_{k\to\infty} \frac{1}{\sqrt{k}} = 0$, we can conclude that it converges!
- Similarly, the series $\displaystyle\sum_{k=3}^{\infty} \frac{(-1)^{k-1}}{\ln k}$ must converge, since the sequence $\left\{\frac{1}{\ln k}\right\}$ is decreasing, with a limit of zero.

**Comment:** You might wonder: if $\lim\limits_{k\to\infty} a_k = 0$, wouldn't we *expect* $\{a_k\}$ to be decreasing? In practice the two properties will usually coincide, but it is possible to construct counterexamples. For example, consider the series

$$1 - \frac{1}{2^2} + \frac{1}{3} - \frac{1}{4^2} + \frac{1}{5} - \frac{1}{6^2} \pm \ldots.$$

Here the sequence of individual terms has a limit of zero, but it is *not* a decreasing sequence! Hence the AST does not apply, and in fact it can be shown that this series diverges.

One nice thing about alternating series is that we have an extremely easy way to approximate their sums:

**The Alternating Series Estimation Theorem (the "ASET")**

| Consider a convergent alternating series $\sum (-1)^k a_k$. If we use the $n^{\text{th}}$ partial sum $s_n$ as an estimate of the sum $s$, then the error satisfies the inequality $\lvert s - s_n \rvert \leq a_{n+1}$. That is, *the truncation error is less than the first term omitted.* |
|---|

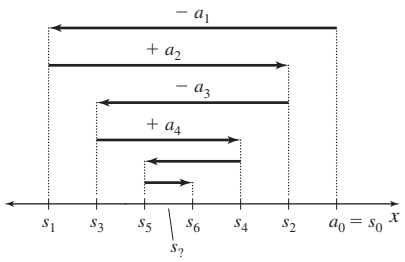
**Example:** Consider the series $\displaystyle\sum_{k=3}^{\infty} (-1)^{k-1} \frac{8}{k^4 \ln k}$. It should be clear that this converges (think about the sequence $\left\{ \dfrac{8}{k^4 \ln k} \right\}$). To estimate its sum, we could calculate

$$s \approx \frac{8}{81 \ln 3} - \frac{8}{256 \ln 4} + \frac{8}{625 \ln 5} - \frac{8}{1296 \ln 6} + \frac{8}{2401 \ln 7} \approx 0.073578296.$$

How accurate is this? Well, the first term we haven't used is $\dfrac{8}{4096 \ln 8} \approx 0.00094 < 0.001$. Therefore we can justifiably claim that

$$\sum_{k=3}^{\infty} (-1)^{k-1} \frac{8}{k^4 \ln k} = 0.074 \pm 0.001.$$

For an idea of why this works (and why the AST works the way it does), consider the following figure:




Because the individual terms alternate, the partial sums move back and forth on the number line. Because they decrease in magnitude, each partial sum gets trapped in between the previous two, and because they have a limit of zero, the partial sums "zero in" on the eventual sum!

**Aside: Rearrangements**

Some alternating series have some strange properties. For example, if you compute the Maclaurin series for $\ln(x+1)$ and evaluate it at $x = 1$, you'll find that

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \frac{1}{11} - \frac{1}{12} + \frac{1}{13} - \frac{1}{14} \ldots$$

(this is $\displaystyle\sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$, the alternating version of the harmonic series).

We're going to do a couple of simple manipulations here, which will lead us to a surprising result. First, we'll just insert some zeroes (this is just to get the terms in an alignment we want):

$$\ln 2 = 0 + 1 + 0 - \frac{1}{2} + 0 + \frac{1}{3} + 0 - \frac{1}{4} + 0 + \frac{1}{5} + 0 - \frac{1}{6} + 0 + \frac{1}{7} \ldots$$

Next, divide the entire series by 2:

$$\frac{\ln 2}{2} = 0 + \frac{1}{2} + 0 - \frac{1}{4} + 0 + \frac{1}{6} + 0 - \frac{1}{8} + 0 + \frac{1}{10} + 0 - \frac{1}{12} + 0 + \frac{1}{14} \ldots$$

Finally, add this expression, term by term, to the original. This gives

$$\frac{3}{2} \ln 2 = 1 + 0 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + 0 + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + 0 + \frac{1}{11} - \frac{1}{6} + \frac{1}{13} + 0 + \ldots$$

We can remove all the zeroes, of course, and so what we have found is that

$$\frac{3}{2} \ln 2 = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \frac{1}{13} + \ldots.$$

Now, take a close look at that result. The series is composed of exactly the same terms as our original series, *but the sum has changed*!

**???**

You might be thinking that we're trying to trick you here... that maybe this is another one of those calculations like Guido Grandi's proof of the existence of God. However, *there's nothing wrong with what we've done.* The sum of an infinite series can indeed change if we change the order of the terms!

How can this be? Well, a first clue is that our original series was strictly alternating; every second term was negative. In our re-arranged series, every *third* term is negative, so it appears that there are now twice as many positive terms as negative terms (and the new sum is indeed larger than the original). Secondly, recall the warning we gave with the definition of infinite series: a *series* is NOT exactly an ordinary sum! Rather, it is a limit of ordinary sums, and the *sum of a series* is defined as the limit of the sequence of partial sums. Well, if we change the order of the terms, *we change the sequence of partial sums, and so it's entirely possible that a different limit may result.*

Having said that, it still seems like an undesirable property. The good news is that not all series behave like this; in fact it's somewhat unusual. The following definitions distinguish between series whose sums are dependent on the order of the terms, and those whose sums are fixed.

**Absolute Convergence vs. Conditional Convergence**

**DEFINITION**: A series $\sum a_k$ is said to be *absolutely convergent* if the series $\sum |a_k|$ is convergent.

**Example:** The series $\sum_{k=1}^{\infty} (-1)^k \frac{1}{k^2}$ is absolutely convergent, since we know that the series $\sum_{k=1}^{\infty} \frac{1}{k^2}$ is convergent. Furthermore, there is a theorem which guarantees that an absolutely convergent series is indeed also convergent in the ordinary sense (absolute convergence can be thought of as a "kind" of convergence), so we do not need to apply the AST here; the series converges.

**DEFINITION**: A series $\sum a_k$ is said to be *conditionally convergent* if it is convergent, but the series $\sum |a_k|$ is divergent.

The classic example of a conditionally convergent series is the one we used above: $\sum (-1)^k \frac{1}{k}$. We know that this series converges (by the AST), but the harmonic series $\sum \frac{1}{k}$ does not.

Now that we have these definitions, we are in a better position to understand the phenomenon we observed above. The sum of an absolutely convergent series will not change if we rearrange the terms. To see why, consider the most extreme situation: suppose we try to take ALL of the positive terms and place them in front of ALL of the negative terms. We

can, in fact, do this with an absolutely convergent series, because the subseries of positive and negative terms will themselves be convergent. That is, we can write

$$\sum_{k=1}^{\infty} (-1)^{k-1} a_k = \sum_{k \text{ even}} a_k \; - \; \sum_{k \text{ odd}} a_k \tag{39}$$

and we see that even in this most extreme rearrangement, the value of the sum must be unchanged. On the other hand, if we consider a conditionally convergent series, then 39 is not a valid calculation, because the two series on the right will be divergent.

Now, let's go back to the series $\displaystyle\sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$ one last time. We can actually force the terms in this series to add up to whatever sum we wish, just by selecting the right order for them! How would we do this? Well, suppose we want the sum to be $S$. We simply start picking out positive terms until we reach a partial sum greater than $S$ (which we will be able to do because the series $\displaystyle\sum_{k \text{ odd}} \frac{1}{k}$ is divergent). Then, we add in negative terms until our partial sum drops back below $S$. We can repeat this procedure forever (in theory, at least), and construct a series whose sum is $S$.

This algorithm won't work for an absolutely convergent series, because at some point we'll be in a position where even if we take ALL of the remaining positive (or negative) terms, we won't be able to get back up (or down) to $S$, because all of those remaining tems together have a finite value.

We won't mention re-arrangements again, but you will often see the distinction made between absolute and conditional convergence; in fact it's part of the statement of our last (and most important) theorem:

## 8. The Ratio Test

Suppose $\displaystyle\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = L$.

If $L < 1$ then the series $\sum a_k$ is absolutely convergent.

If $L > 1$ then the series $\sum a_k$ is divergent.

If $L = 1$ then the test fails (the series may be absolutely convergent, conditionally convergent, or divergent).

The fact that the test fails when $L = 1$ is important; if it wasn't for this we'd hardly need the other seven tests!

**Examples:**

- Consider the series $\sum (-1)^k \dfrac{2^k}{k!}$. This happens to be alternating, but we don't have to use the AST; the Ratio Test is easier.

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\left|\frac{2^{k+1}}{(k+1)!} \div \frac{2^k}{k!}\right|$$

$$= \lim_{k\to\infty}\left|\frac{2^{k+1}}{(k+1)!} \cdot \frac{k!}{2^k}\right|$$

$$= \lim_{k\to\infty} \frac{2}{k+1}$$

$$= 0 \ < 1,$$

  so the series converges absolutely.

- Consider the series $\displaystyle\sum_{k=1}^{\infty} \frac{3^k}{k^2 2^k}$. This time we find

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\left|\frac{3^{k+1}}{(k+1)^2\, 2^{k+1}} \cdot \frac{k^2 2^k}{3^k}\right|$$

$$= \lim_{k\to\infty} \frac{3}{2}\left(\frac{k}{k+1}\right)^2 = \frac{3}{2} \ > 1,$$

  so this series is divergent.

One final note: there is actually one more test, called the Root Test. It works exactly like the Ratio Test, but uses the limit $\lim\limits_{k\to\infty} \sqrt[k]{|a_k|}$. The logic behind both tests is that they test whether $\sum a_k$ "behaves like" a geometric series when $k$ becomes large (apply either test to $\sum ar^k$ and you get $L = r$). The Root Test is useful for series in which everything appears raised to the power of $k$, such as $\displaystyle\sum \left(\frac{1}{\tan^{-1} k}\right)^k$, but that's a structure you will rarely encounter.

## 22 Power Series

We now return to our discussion of Taylor series. More generally, a *power series centered at* $x_0$ is any series of the form

$$\sum_{k=0}^{\infty} c_k (x - x_0)^k = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + \dots . \tag{40}$$

**Aside:** There's really no difference between a "Taylor series" and a "power series"; it's just that using the term "Taylor series" implies that you've used knowledge of Taylor's formula in some way to obtain the series, whereas the term "power series" does not. Usually, we'll call our series a Taylor series if we've obtained it from a function, and we'll call it a power series if we're using it to *define* a function.

Given a power series (40), we'd like to know for which values of $x$ it converges. So, we apply the Ratio Test!

We have

$$\begin{aligned}
\lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| &= \lim_{k\to\infty} \left| \frac{c_{k+1} (x - x_0)^{k+1}}{c_k (x - x_0)^k} \right| \\
&= \lim_{k\to\infty} \left| \frac{c_{k+1}}{c_k} \right| \cdot |x - x_0| \\
&= |x - x_0| \lim_{k\to\infty} \left| \frac{c_{k+1}}{c_k} \right| .
\end{aligned}$$

Therefore the series converges absolutely if

$$|x - x_0| \lim_{k\to\infty} \left| \frac{c_{k+1}}{c_k} \right| < 1,$$

i.e. if

$$|x - x_0| < \frac{1}{\lim\limits_{k\to\infty} \left| \frac{c_{k+1}}{c_k} \right|} = \lim_{k\to\infty} \left| \frac{c_k}{c_{k+1}} \right| .$$

Now, let $R = \lim\limits_{k\to\infty} \left| \frac{c_k}{c_{k+1}} \right|$ (assuming that the limit exists). We've then established that the series $\sum\limits_{k=0}^{\infty} c_k (x - x_0)^k$ converges absolutely if $|x - x_0| < R$. To be more precise, we can state the following:

- If $R = 0$, then the series converges *only* at $x = x_0$.

- If $R = \infty$, then the series converges for *all* $x$.
- If $0 < R < \infty$, then the series converges absolutely for $x \in (x_0 - R,\ x_0 + R)$ and diverges for $x < x_0 - R$ and for $x > x_0 + R$. At the two points $x = x_0 \pm R$, the test gives no conclusion; the series may converge absolutely, it may converge conditionally, or it may diverge (now you see why we have been calling $x_0$ the "center" of the series).

We refer to the number $R$ as the *radius of convergence*. We will also speak of the *interval of convergence*; to find this we need to investigate the behaviour of the series at the two endpoins $x = x_0 - R$ and $x = x_0 + R$. This is the only purpose for which we really need our other convergence tests!

**Comment:** Instead of learning the above formula for $R$, it may be a good idea just to apply the Ratio Test in each instance. This will give you one less formula to remember, and it may help you if you encounter series appearing in non-standard forms. If you want to use the formula, think of it as a *shortcut* for the Ratio Test for use on series of the form $\sum c_k (x - x_0)^k$. Be careful if you're presented with a series with a different form (something like $\sum c_k (x - x_0)^{2k}$, for example). When in doubt, go back to the Ratio Test.

**Examples:**

a) Consider the series $\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$. For what values of $x$ does this converge?

Well, apply the Ratio Test:

$$\lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| = \lim_{k\to\infty} \left| \frac{x^{k+1}}{(k+1)!} \cdot \frac{k!}{x^k} \right|$$

$$= |x| \lim_{k\to\infty} \frac{1}{k+1}.$$

This limit is zero for all $x$, so the series converges *for all* $x$.

Note: You might recall that we claimed earlier that the sequence $\left\{ \dfrac{x^k}{k!} \right\}$ has a limit of zero as $k \to \infty$, for any $x$. We've just proved[26] that result! (Do you understand how?)

Also, we've seen this particular series before; do you recognize it?

[26] Well, almost. We didn't actually prove the convergence tests!

b) Consider the series $\sum_{k=0}^{\infty} k!\, x^k$. This time the Ratio Test gives us

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\left|\frac{(k+1)!\, x^{k+1}}{k!\, x^k}\right| = \lim_{k\to\infty} k\,|x|\,.$$

This limit fails to exist (the sequence approaches infinity) unless $x = 0$, and therefore the series converges only at its center (which is not very useful).

c) Consider the series $\sum_{k=1}^{\infty} \frac{(x-3)^k}{k\ 4^k}$. For this one we find

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\left|\frac{(x-3)^{k+1}}{(k+1)\ 4^{k+1}}\cdot\frac{k\ 4^k}{(x-3)^k}\right|$$

$$= \lim_{k\to\infty}\frac{k}{4\,(k+1)}\,(x-3) = \frac{1}{4}\,|x-3|\,.$$

We can conclude that the series converges absolutely if $\frac{1}{4}\,|x-3| < 1$, that is if $|x-3| < 4$. Therefore the radius of convergence is 4. To find the interval of convergence we need to test those two points where $|x-3| = 4$ exactly. Well, if $x - 3 = 4$ (that is, if $x = 7$), then our series is $\sum_{k=1}^{\infty} \frac{1}{k}$, which we know to be divergent. Meanwhile, if $x - 3 = -4$ (that is, if $x = -1$), then the series is $\sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$, which we know to be convergent. Therefore the interval of convergence for this series is the interval $[-1, 7)$.

Note: The status of the endpoints is of mathematical interest, but it will rarely be important in applications. Even when we find the series to be convergent at an endpoint, it will usually converge extremely slowly (the series in the above example happens to converge to $-\ln 2$, but it would take thousands of terms to give you an approximation of any useful accuracy).

## Manipulation of Power Series

You've already seen some of the tricks we can employ to calculate Taylor polynomials without having to calculate multiple derivatives of the original function. Fortunately, we can use these same tricks to calculate entire Taylor series, and we don't have to re-calculate the radius of convergence!

**Theorem:**

If the series $\sum c_k (x - x_0)^k$ has radius of convergence $R$, then we can

- differentiate it (term-by-term)
- integrate it (term-by-term)
- multiply through by a constant (term-by-term)
- add it (term-by-term) to another series of radius of convergence $\geq R$

and the result will also have radius of convergence $R$.

Furthermore, since Taylor series are unique for each $f$, $x_0$, and $n$, if we perform these operations on Taylor series the results will be the Taylor series for the differentiated / integrated / multiplied / summed functions.

In theory this theorem can be extended to multiplication or division of two series, but there is a complication; those operations are not carried out term-by-term! This means that we'll rarely be able to find the full series. However, the theorem still tells us something about the interval over which we can expect Taylor polynomials obtained in this way to be useful (we'll do some examples of these sorts of calculations shortly).

**Example:** Here's an example of the sorts of things we can do. Consider the series $\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + x^4 + \ldots$. This is a geometric series with common ratio $x$, so we know its sum: it's $\frac{1}{1-x}$, and it converges for $|x| < 1$ (so its radius of convergence is 1). That is,

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + x^4 + \ldots$$

We could differentiate this, and thus determine that

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} k x^{k-1} = 1 + 2x + 3x^2 + 4x^3 + \ldots.$$

This is a much easier way to determine the Maclaurin series for $\frac{1}{(1-x)^2}$ than using the standard formula, and we don't even have to apply the ratio test to determine the radius of convergence; we know it can't have changed when we differentiated, so it's 1.

**Comment:** Notice that we've written the new series with the index starting at $k = 1$, even though the original series started at $k = 0$. This is because the first term was a constant, so it disappeared when we differentiated. This doesn't always happen, because the first term won't always be constant! The point is that when differentiating a series, you should always ask yourself if there are any constant terms being eliminated.

Continuing with our example, we could also *integrate* the original series. If we do this, we find that

$$-\ln(1-x) = C + \sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1}. \tag{41}$$

What do we do with the constant of integration? Well, there's one value at which we can evaluate the series easily, namely $x = 0$. Plugging this into Equation 41, we immediately see that $C = 0$. Therefore, multiplying through by -1, we have

$$\ln(1-x) = -\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \frac{1}{4}x^4 - \dots.$$

This is the Maclaurin series for $\ln(1-x)$, and this *also* has $R = 1$.

We could also *add* these two results:

$$\begin{aligned}
\frac{1}{(1-x)^2} + \ln(1-x) &= \sum_{k=1}^{\infty} kx^{k-1} - \sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} \\
&= \sum_{n=0}^{\infty} (n+1)\, x^n - \sum_{n=1}^{\infty} \frac{x^n}{n} && * \\
&= 1 + \sum_{n=1}^{\infty} (n+1)\, x^n - \sum_{n=1}^{\infty} \frac{x^n}{n} && ** \\
&= 1 + \sum_{n=1}^{\infty} \left[ n + 1 - \frac{1}{n} \right] x^n \\
&= 1 + x + \frac{5}{2}x^2 + \frac{11}{3}x^3 + \frac{19}{4}x^4 + \dots.
\end{aligned}$$

and again, we know that $R = 1$ for this series as well.

* Here we've re-indexed the series, using $n = k - 1$ in the first series, and $n = k + 1$ in the second series. We do this so that we can add the two series together in summation notation; this requires us to have $x$ raised to the same power in both.

** Similarly, in order to add the two series together in summation notation we must have the same starting value for the index. To achieve this, we've removed the $n = 0$ term from the first sum and written it as its own term.

**Note:** Our theorem concerns the *radius* of convergence only; it does not guarantee that the *interval* will be unchanged. In fact, we may lose convergence at an endpoint when we differentiate, and we may gain convergence at and endpoint when we integrate. If you look at our results above, you'll see that that's exactly what happened; the original (geometric) series diverges at both 1 and -1, but the series for $\ln(1-x)$ actually converges at -1.

**Example:** The geometric series is an extremely useful one. We can use it to obtain Taylor expansions for the logarithm function (as we just did), the arctangent (by making the substitution $x \to -t^2$ and *then* integrating), and a wide variety of rational functions. For example, suppose we want the Maclaurin series for $f(x) = \dfrac{x}{3+2x}$. We can find this as follows:

$$\begin{aligned}
\frac{x}{3+2x} &= x\left(\frac{1}{3+2x}\right) \\
&= \frac{x}{3}\cdot\left(\frac{1}{1+\frac{2}{3}x}\right) \\
&= \frac{x}{3}\cdot\left(\frac{1}{1-\left(-\frac{2}{3}x\right)}\right) \\
&= \frac{x}{3}\cdot\sum_{k=0}^{\infty}\left(-\frac{2}{3}x\right)^k \qquad \dagger \\
&= \sum_{k=0}^{\infty}(-1)^k\,\frac{2^k x^{k+1}}{3^{k+1}} \qquad \ddagger
\end{aligned}$$

What is the radius of convergence? Well, at the step marked † we knw the series converges if $\left|-\dfrac{2}{3}x\right| < 1$, that is if $|x| < \frac{3}{2}$. At the following step, marked ‡ , we know that factoring in the $\frac{1}{3}$ doesn't affect the radius of convergence. As for the extra factor of $x$, we can view this as a multiplication of two series. The first series simply happens to have only one non-zero term (it's a finite series), so it has radius of convergence $R = \infty$. Because it has just the one term we are in fact able to compute the entire series resulting from the multiplication, and we know the smaler radius of convergence is the one that matters. Therefore our series for $f(x) = \dfrac{x}{3+2x}$ has radius $R = \frac{3}{2}$.

### The Basic Building Blocks

With tricks like those we've just discussed, we can find Maclaurin series for most of the functions we'll encounter from just a handful of basic building blocks. Specifically, you'll find it useful to remember the following:

$$\begin{aligned}
\frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k, && \text{for } |x|<1 \\
e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!}, && \text{for all } x \\
\sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, && \text{for all } x \\
\cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, && \text{for all } x
\end{aligned}$$

In addition, the Binomial Series provides a quick way of writing down the Maclaurin series for functions of a certain type. This is a generalization of the Binomial Theorem which you may have seen in high school; whereas the Binomial Theorem allows us to expand expressions of the form $(a+b)^m$ easily, the Binomial Series removes the restriction that $m$ need be an integer (and, to make it easier to use with functions, we set $a=1$). Specifically, it allows us to write

$$(1+x)^m = 1 + mx + m(m-1)\frac{x^2}{2!} + m(m-1)(m-2)\frac{x^3}{3!} + \dots$$
$$\dots + m(m-1)(m-2)\cdots(m-n+1)\frac{x^n}{n!} + \dots.$$

Furthermore, it has radius of convergence 1.

**Example:** To find the first few terms of the Maclaurin series for the function $\dfrac{1}{\sqrt{2-x}}$, we can write

$$\begin{aligned}
\frac{1}{\sqrt{2-x}} &= \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{1-\frac{x}{2}}}\right) \\
&= \frac{1}{\sqrt{2}}\left(1-\frac{x}{2}\right)^{-\frac{1}{2}} \\
&= \frac{1}{\sqrt{2}}\left[1 + \left(-\frac{1}{2}\right)\left(-\frac{x}{2}\right) + \frac{1}{2!}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{x}{2}\right)^2 + \right. \\
&\qquad \left. + \frac{1}{3!}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)\left(-\frac{x}{2}\right)^3 + \dots\right] \\
&= \frac{1}{\sqrt{2}}\left[1 + \frac{1}{4}x + \frac{3}{32}x^2 + \frac{5}{128}x^3 + \dots\right]
\end{aligned}$$

and we know that this series converges for $\left|-\dfrac{x}{2}\right| < 1$, i.e. for $|x|<2$.

## 23 The "Big-O" Order Symbol

To motivate the concept we're about to introduce, consider the following situation: suppose you've shown that the $4^{\text{th}}$-order Maclaurin polynomial for a certain function is, say, $1 - x^3 + \frac{1}{10}x^4$. Suppose also that you go ahead and calculate that the $5^{\text{th}}$ derivative of this function is zero when $x$ is zero, so that this same expression is *also the* $5^{th}$*-order Maclaurin polynomial*. How would you incorporate this extra piece of information in your work? Well, in the notation we've used in this course we have a fairly efficient way to do this; we could state that $f(x) = 1 - x^3 + \frac{1}{10}x^4 + R_5(x)$. Even if we don't feel the need to proceed with finding an upper bound on the magnitude of the error term, the fact that we've written $R_5$ shows clearly that we know the $x^5$ term to be zero. What we are about to introduce is essentially an alternative to that notation which is universally recognized. The difficulty is that it is an application of a broader concept which is rather more difficult to grasp.

**Definition:**

Given two functions $f$ and $g$, we say that "$f$ is of order $g$ as $x \to x_0$" and write

$$f(x) = \mathcal{O}(g(x)) \text{ as } x \to x_0$$

if there exists a constant $A$ (greater than zero) such that

$$|f(x)| \leq A\,|g(x)|$$

on some interval around $x_0$ (although the point $x_0$ itself may be excluded from the interval, since the idea is to describe the behaviour of $f$ in the *limit* as we approach $x_0$).

For our purposes the function $g$ will be a power of $(x - x_0)$, and of course if we're dealing with Maclaurin series then it will simply be a power of $x$.

**Note:** You might be familiar with a similar concept used in computer science, which deals with behaviour as $n \to \infty$ instead of $x \to x_0$.

**Examples:**

1. Since we know that $\left|x^3\right| \leq \left|x^2\right|$ for all $x \in [-1, 1]$, we can state that

$$x^3 = \mathcal{O}\left(x^2\right) \text{ as } x \to 0.$$

   Since it is also true that $\left|x^3\right| \leq |x|$ on the same interval, we can also state that

$$x^3 = \mathcal{O}\left(x\right) \text{ as } x \to 0.$$

   Both statements are correct (just look at our definition; in these examples the constant $A$ is just 1). In fact, we can similarly state that $x^3 = \mathcal{O}\left(x^3\right)$ as $x \to 0$, and $x^3 = \mathcal{O}\left(1\right)$ as $x \to 0$ (this last statement just says that $x^3$ is *bounded* near zero; $\left|x^3\right| \leq A$ for some number $A$ when $x$ is small enough).

2. On the other hand, it is *not* true that $x^3 = \mathcal{O}\left(x^4\right)$ as $x \to 0$. We *could* say instead that $x^3 = \mathcal{O}\left(x^4\right)$ as $x \to \infty$, since $\left|x^3\right| \leq \left|x^4\right|$ when $|x| \geq 1$. The "as $x \to x_0$" part of the definition is important!

3. A particularly well-known inequality states that $|\sin x| \leq |x|$, for all $x$. This allows us to state that

$$\sin x = \mathcal{O}\left(x\right) \text{ as } x \to 0.$$

   We can also rearrange the inequality as $\left|\dfrac{\sin x}{x}\right| \leq 1$, for all $x \neq 0$. Hence we may write

$$\frac{\sin x}{x} = \mathcal{O}\left(1\right) \text{ as } x \to 0.$$

   This is a convenient ways of stating that even though the function $\dfrac{\sin x}{x}$ isn't defined at 0, it is bounded nearby (that is, it does *not* have a vertical asymptote there).

4. Consider the function $\dfrac{x^2}{8 + x^3}$, on the interval $[-1, 1]$. On this interval, we can say that $\left|\dfrac{x^2}{8 + x^3}\right| = \dfrac{1}{|8 + x^3|}\left|x^2\right| \leq \dfrac{1}{7}\left|x^2\right|$, and therefore

$$\frac{x^2}{8 + x^3} = \mathcal{O}\left(x^2\right) \text{ as } x \to 0.$$

**Comments:**

- Although this notation is standard, it is really a grievous misuse of the "=" sign. We've said that $x^3 = \mathcal{O}(x)$ as $x \to 0$, and we've said that $\sin x = \mathcal{O}(x)$ as $x \to 0$, but of course we are *not* claiming that $x^3 = \sin x$! One could argue that "$\in$" would be a more appropriate symbol, since what we really mean is that $x^3$ and $\sin x$ both belong to a certain *class* of functions.

- It is important to appreciate that the definition of the $\mathcal{O}$ symbol assigns no importance at all to the *value* of the constant $A$. So, for example, we can write

$$\sin x = \mathcal{O}(x) \text{ as } x \to 0,$$

$$\text{or} \qquad 10 \sin x = \mathcal{O}(x) \text{ as } x \to 0,$$

$$\text{or even} \qquad 10^{10} \sin x = \mathcal{O}(x) \text{ as } x \to 0.$$

  The functions $\sin x$, $10 \sin x$, and $10^{10} \sin x$ all fall into the same category. In fact, that's really the point; we're trying to say something about how the values of the function change as $x$ approaches $x_0$, without saying anything at all about how large those values are. That is, the "order" symbol has nothing to do with order of magnitude. Instead, as we're about to see, it has more to do with the order of the Taylor polynomials we can use as approximations to the function.

### Using the Order Symbol in Conjunction with Taylor's Inequality

If you're wondering what the point of all this is, consider the statement of Taylor's Inequality: if $R_n(x)$ is the error associated with using the $n^{\text{th}}$-order Taylor polynomial (centered at $x_0$) as an approximation to $f(x)$, then

$$|R_n(x)| \leq \frac{K |x - x_0|^{n+1}}{(n+1)!},$$

where $\left|f^{(n+1)}(t)\right| \leq K$ for all $t$ between $x_0$ and $x$. Now, compare this expression to our definition of the "Big-$\mathcal{O}$"; it says precisely that

$$R_n(x) = \mathcal{O}\left((x-x_0)^{n+1}\right) \text{ as } x \to x_0!$$

Now you can see the advantage of the notation. In essence it gives us a convenient way to say "I haven't calculated the constant $K$ which is required in Taylor's Inequality, but I can tell you the power of $(x-x_0)$ that appears."

In practice, we rewrite the above expression slightly. From the way we originally defined $R_n(x)$, we can write

$$f(x) - P_{n,x_0}(x) = \mathcal{O}\left((x-x_0)^{n+1}\right) \text{ as } x \to x_0.$$

From this we obtain the expression $f(x) = P_{n,x_0}(x) + \mathcal{O}\left((x-x_0)^{n+1}\right)$ as $x \to x_0$. Strictly speaking, we shouldn't be allowed to do this; it only *looks* as though we can move the terms around this way because of the way we're abusing the "=" sign! However, it happens to be convenient; if we allow ourselves this little bit of freedom we have a nice, simple way of tracking the order of the error terms when we're manipulating Taylor polynomials. Here's how this works:

**Examples:**

a) Since $\sqrt{1+x} = 1 + \frac{1}{2}x + \mathcal{O}\left(x^2\right)$ as $x \to 0$, and $\sin x = x + \mathcal{O}\left(x^3\right)$ as $x \to 0$, we know immediately that

$$\sqrt{1+x} + \sin x = 1 + \frac{3}{2}x + \mathcal{O}\left(x^2\right) + \mathcal{O}\left(x^3\right) \text{ as } x \to 0.$$

In fact we can simplify this slightly; anything which is of order $x^3$ is also of order $x^2$ (as $x \to 0$, of course), so we may write

$$\sqrt{1+x} + \sin x = 1 + \frac{3}{2}x + \mathcal{O}\left(x^2\right) \text{ as } x \to 0.$$

b) For the hyperbolic sine function, we may write

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

$$= \frac{1}{2}\left[\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \mathcal{O}\left(x^5\right)\right)\right.$$

$$\left. - \left(1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} + \mathcal{O}\left(x^5\right)\right)\right].$$

Now, an unknown function of order $x^5$ minus an unknown function of order $x^5$ is just another unknown function of order $x^5$, so after simplification we have

$$\sinh x = x + \frac{x^3}{3!} + \mathcal{O}\left(x^5\right), \text{ as } x \to 0.$$

Comment: This is an elegant result. Recall from Math 117 that the hyperbolic sine function was originally defined as *the odd component of the exponential function* (and the hyperbolic cosine is the even component). Since the Maclaurin series for the exponential function is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots,$$

it should really be no surprise that the Maclaurin series for the hyperbolic functions are

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \ldots,$$

and

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!}.$$

From these expressions you can also see why these function share so many properties with the trigonometric functions!

c) Substitution works exactly as you might expect. Since

$$\sqrt{1+x} = 1 + \frac{x}{2} + \mathcal{O}\left(x^2\right), \text{ as } x \to 0,$$

we can, for example, state immediately that

$$\sqrt{1+u^4} = 1 + \frac{u^4}{2} + \mathcal{O}\left(u^8\right), \text{ as } u \to 0.$$

d) Multiplication of Taylor series now becomes practical. We won't often be able to find the entire series that we're looking for, but we can keep track of how many terms we do indeed know. For example, we can write

$$\begin{aligned} e^x \sin x &= \left[1 + x + \frac{x^2}{2} + \mathcal{O}\left(x^3\right)\right]\left[x + \mathcal{O}\left(x^3\right)\right] \\ &= x + x^2 + \frac{x^3}{2} + x\mathcal{O}\left(x^3\right) \\ &+\mathcal{O}\left(x^3\right) + x\mathcal{O}\left(x^3\right) + \frac{x^2}{2}\mathcal{O}\left(x^3\right) + \left[\mathcal{O}\left(x^3\right)\right]^2, \qquad \text{as } x \to 0. \end{aligned}$$

Now, a little bit of thought should reveal that all of these unknown terms are of order $x^3$ (because anything which is of order $x^4$, $x^5$, or $x^6$ is *also* of order $x^3$!). Furthermore, since the expansion contains unknown terms of order $x^3$, there's really no point in retaining the $\frac{x^3}{2}$ we've found (it isn't telling us anything about the coefficient of $x^3$ in our Taylor series, because $\frac{1}{2}$ plus an unknown number is just another unknown number). With what we've done, we can conclude that

$$e^x \sin x = x + x^2 + \mathcal{O}\left(x^3\right), \text{ as } x \to 0.$$

e) Division is still hard, but it can be done. For example,

$$\tan x = \frac{\sin x}{\cos x} = \frac{x - \frac{x^3}{6} + \frac{x^5}{120} + \mathcal{O}\left(x^7\right)}{1 - \frac{x^2}{2} + \frac{x^4}{24} + \mathcal{O}\left(x^6\right)}.$$

We can tackle this with long division, like this:

$$
\begin{array}{r|l}
 & x+\frac{1}{3}x^3+\frac{2}{15}x^5+\mathcal{O}\left(x^7\right) \\
\hline
1-\frac{x^2}{2}+\frac{x^4}{24}+\mathcal{O}\left(x^6\right) & x-\frac{1}{6}x^3+\frac{1}{120}x^5+\mathcal{O}\left(x^7\right) \\
 & x-\frac{1}{2}x^3+\frac{1}{24}x^5+\mathcal{O}\left(x^7\right) \\
\hline
 & \frac{1}{3}x^3-\frac{1}{30}x^5+\mathcal{O}\left(x^7\right) \\
 & \frac{1}{3}x^3-\frac{1}{6}x^5+\mathcal{O}\left(x^7\right) \\
\hline
 & \frac{2}{15}x^5+\mathcal{O}\left(x^7\right) \\
 & \frac{2}{15}x^5+\mathcal{O}\left(x^7\right) \\
\hline
 & \mathcal{O}\left(x^7\right)
\end{array}
$$

We conclude that

$$\tan x = x+\frac{1}{3}x^3+\frac{2}{15}x^5+\mathcal{O}\left(x^7\right) \text{ as } x\to 0.$$

If we think of the order symbol as a placeholder for all of the omitted terms in our Taylor series, then the calculations in the above examples should be clear (although you might need some guidance with the long division procedure). If you wish, though, you can think in terms of the following rules for the "algebra of the Big-$\mathcal{O}$":

All of the following hold as $x\to 0$:

1. $k\mathcal{O}\left(x^n\right)=\mathcal{O}\left(x^n\right)$, for any constant $k$.
2. $\mathcal{O}\left(x^m\right)+\mathcal{O}\left(x^n\right)=\mathcal{O}\left(x^q\right)$, where $q$ is the lesser of $m$ and $n$.
3. $\mathcal{O}\left(x^m\right)\cdot\mathcal{O}\left(x^n\right)=\mathcal{O}\left(x^{m+n}\right)$.
4. $\left[\mathcal{O}\left(x^n\right)\right]^m=\mathcal{O}\left(x^{mn}\right)$.
5. $\frac{\mathcal{O}(x^m)}{x^n}=\mathcal{O}\left(x^{m-n}\right)$.

Note that we cannot state a general rule for simplifying $\dfrac{\mathcal{O}\left(x^m\right)}{\mathcal{O}\left(x^n\right)}$, because in that expression we cannot tell what the lowest power in the denominator is (recall our very first example; an expression such as $\mathcal{O}\left(x^2\right)$ could actually stand for $x^3$, or $x^4$, etc).

### Evaluation of Limits using Taylor Polynomials and the Order Symbol

Since everything we've discussed applies "as $x \to x_0$", it lends itself neatly to the evaluation of certain kinds of limits. Here's a trivial example to demonstrate the idea:

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{x + \mathcal{O}\left(x^3\right)}{x}$$

$$= \lim_{x\to 0} \left[1 + \mathcal{O}\left(x^2\right)\right]$$

$$= 1.$$

Of course, we've known this limit for a long time, but you can see the principle. This technique is going to work especially well if we have a simple power of $x$ in the denominator. In fact, it's occasionally the easiest method available to us; to see that, consider the following example, and try using L'Hôpital's Rule instead!

$$\begin{aligned}
\lim_{x\to 0} \frac{x\cos x - \sin x}{x^3} &= \lim_{x\to 0} \frac{x\left[1 - \frac{x^2}{2!} + \mathcal{O}\left(x^4\right)\right] - \left[x - \frac{x^3}{3!} + \mathcal{O}\left(x^5\right)\right]}{x^3} \\
&= \lim_{x\to 0} \frac{\left(x - \frac{x^3}{2} + \mathcal{O}\left(x^5\right)\right) - \left(x - \frac{1}{6}x^3 + \mathcal{O}\left(x^5\right)\right)}{x^3} \\
&= \lim_{x\to 0} \frac{-\frac{1}{3}x^3 + \mathcal{O}\left(x^5\right)}{x^3} \\
&= \lim_{x\to 0} -\frac{1}{3} + \mathcal{O}\left(x^2\right) \\
&= -\frac{1}{3}.
\end{aligned}$$

## 24 Taylor Series: the Two-Variable Case

We spent the first half of this course extending the concepts of single-variable calculus to functions of more than one variable, and now we have spent the second half back in the world of single-variable functions. So, can we extend our newest concepts to multivariate functions as well? Yes, of course we can!

To determine what a Taylor expansion of a function of two variables should look like, let's start with what we know about the single-variable variety. Consider a function $f(x, y)$, and

suppose that we want an approximation to this which is valid for values of $x$ and $y$ near the point $(x_0, y_0)$. If we think of one variable, say $y$, as being fixed, then $f(x, y)$ becomes a function of $x$ alone, and we can expand it in a Taylor series centered at $x_0$ (the only difference being that the derivatives of $f$ need to be expressed as *partial* derivatives):

$$f(x,y) = f(x_0,y) + f_x(x_0,y)\cdot(x-x_0) + \frac{1}{2!}f_{xx}(x_0,y)\cdot(x-x_0)^2 + \dots \tag{42}$$

Now, the terms $f(x_0, y)$, $f_x(x_0, y)$, $f_{xx}(x_0, y)$, etc. are, of course, not truly constant, but depend on the second variable $y$. Therefore we should be able to expand *each* of them in a Taylor series centered at $y_0$:

$$f(x_0,y) = f(x_0,y_0) + f_y(x_0,y_0)\cdot(y-y_0) + \frac{1}{2!}f_{yy}(x_0,y_0)\cdot(y-y_0)^2 + \dots \tag{43}$$

$$f_x(x_0,y) = f_x(x_0,y_0) + f_{xy}(x_0,y_0)\cdot(y-y_0) + \frac{1}{2!}f_{xyy}(x_0,y_0)\cdot(y-y_0)^2 + \dots \tag{44}$$

$$f_{xx}(x_0,y) = f_{xx}(x_0,y_0) + f_{xxy}(x_0,y_0)\cdot(y-y_0) + \frac{1}{2!}f_{xxyy}(x_0,y_0)\cdot(y-y_0)^2 + \dots \tag{45}$$

etc.

Substituting each of these into Equation 42 produces an infinite series *of infinite series!* This sounds unwieldy, but it's really just one series; all we need to do is organize the terms in a sensible way. We are trying to construct a Taylor series, from which we should be able to extract a linear approximation, a quadratic approximation, and so on. If we look at the above equations carefully, we see that there is only one constant term: $f(x_0, y_0)$. There are two linear terms: $f_x(x_0, y_0)\cdot(x - x_0)$ and $f_y(x_0, y_0)\cdot(y - y_0)$. There are then three quadratic terms, four quintic terms, and so on (these might be easier to identify if you realize that, as in the single variable case, the linear terms will involve first-order derivatives, the quadratic terms will involve second-order derivatives, etc; these show up along the diagonals in the array

of Equations 43, 44, and 45). We can write the Taylor series out as follows:

$$\begin{aligned} f(x,y) = \quad & f(x_0,y_0) \\ & + f_x(x_0,y_0)\cdot(x-x_0) + f_y(x_0,y_0)\cdot(y-y_0) \\ & + \tfrac{1}{2!}f_{xx}(x_0,y_0)\cdot(x-x_0)^2 + f_{xy}(x_0,y_0)\cdot(x-x_0)(y-y_0) + \tfrac{1}{2!}f_{yy}(x_0,y_0)\cdot(y-y_0)^2 \\ & + \ldots \end{aligned} \tag{46}$$

The first two lines here should look familiar: $f(x,y) \approx f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$ is the linear approximation from section 3! Reading on from there, the first three lines give the quadratic approximation, and so on. Our standard notation is getting slightly tedious, so let's introduce an abbreviated form. Let $P_0 = (x_0, y_0)$, let $h = x - x_0$, and let $k = y - y_0$. Then Equation 46 becomes (after including more terms and factoring out all the factorials)

$$\begin{aligned} f(x,y) = \quad & f(P_0) \\ & + f_x(P_0)h + f_y(P_0)k \\ & + \tfrac{1}{2!}\left[f_{xx}(P_0)h^2 + 2f_{xy}(P_0)hk + f_{yy}(P_0)k^2\right] \\ & + \tfrac{1}{3!}\left[f_{xxx}(P_0)h^3 + 3f_{xxy}(P_0)h^2k + 3f_{xyy}(P_0)hk^2 + f_{yyy}(P_0)k^3\right] \\ & + \ldots \end{aligned} \tag{47}$$

The pattern should now be clear; much of the structure of the single-variable Taylor series formula is retained, but *all* of the partial derivatives must appear, with the coefficients coming from Pascal's Triangle. [27]

**Example:** Find the 3rd-order Taylor approximation to the function $f(x,y) = e^x \cos y$, centered at $(0,0)$.

[27] There's a nice explanation for this: the term $2f_{xy}(P_0)hk$ actually represents the *two* terms $f_{xy}(P_0)hk$ and $f_{yx}(P_0)hk$, which happen to be equal (as long as all of the derivatives are continuous, which is a condition we need for the Taylor series to exist, anyway). Similarly, the $f_{xxy}$ term is actually composed of the three terms containing $f_{xxy}$, $f_{xyx}$, and $f_{yxx}$, and so on!

**Solution:**

$$\begin{aligned}
f(x,y) = e^x \cos y \quad &\Longrightarrow \quad f(0,0) = 1 \\
f_x(x,y) = e^x \cos y \quad &\Longrightarrow \quad f_x(0,0) = 1 \\
f_y(x,y) = -e^x \sin y \quad &\Longrightarrow \quad f_y(0,0) = 0 \\
f_{xx}(x,y) = e^x \cos y \quad &\Longrightarrow \quad f_{xx}(0,0) = 1 \\
f_{xy}(x,y) = -e^x \sin y \quad &\Longrightarrow \quad f_{xy}(0,0) = 0 \\
f_{yy}(x,y) = -e^x \cos y \quad &\Longrightarrow \quad f_{yy}(0,0) = -1 \\
f_{xxx}(x,y) = e^x \cos y \quad &\Longrightarrow \quad f_{xxx}(0,0) = 1 \\
f_{xxy}(x,y) = -e^x \sin y \quad &\Longrightarrow \quad f_{xxy}(0,0) = 0 \\
f_{xyy}(x,y) = -e^x \cos y \quad &\Longrightarrow \quad f_{xyy}(0,0) = -1 \\
f_{yyy}(x,y) = e^x \sin y \quad &\Longrightarrow \quad f_{yyy}(0,0) = 0
\end{aligned}$$

Substituting all of these values into our formula (Equation 47), with $P_0 = (0,0)$, $h = x - x_0 = x$, and $k = y - y_0 = y$, we find that

$$e^x \cos y \approx 1 + x + \frac{x^2}{2} - \frac{y^2}{2} + \frac{x^3}{6} - \frac{xy^2}{2}.$$

Having done one example using the formula directly, we should point out that, as in the single-variable case, we can often find an easier way. Since we already know the Maclaurin series for both the exponential and the cosine, we could have proceeded this way:

$$\begin{aligned}
e^x \cos y &= \left[1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \mathcal{O}(x^4)\right]\left[1 - \frac{y^2}{2} + \mathcal{O}(y^4)\right] \\
&= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \mathcal{O}(x^4) - \frac{y^2}{2} - \frac{xy^2}{2} - \frac{x^2y^2}{4} - \frac{x^3y^2}{12} - \frac{y^2}{2}\mathcal{O}(x^4) \\
&\qquad + \mathcal{O}(x^4)\mathcal{O}(y^4) + \mathcal{O}(y^4)\left[1 + x + \frac{x^2}{2} + \ldots\right] \\
&= 1 + x + \frac{x^2}{2} - \frac{y^2}{2} + \frac{x^3}{6} - \frac{xy^2}{2} + \mathcal{O}(x^4) + \mathcal{O}(x^2y^2) + \mathcal{O}(y^4)
\end{aligned}$$

Here we've dropped the $x^2y^2$ and $x^3y^2$ terms because they don't belong in the third-order approximation (they are $4^{\text{th}}$ and $5^{\text{th}}$ order, respectively, and technically they are both $\mathcal{O}(x^2y^2)$). The Big-$\mathcal{O}$ notation is a little bit cumbersome here, and you can see that it gets worse as we include more terms. Because of this, it's common to make the assumption that $y$ is of order

$x$. This allows us to write our result as

$$e^x \cos y = 1 + x + \frac{x^2}{2} - \frac{y^2}{2} + \frac{x^3}{6} - \frac{xy^2}{2} + \mathcal{O}(x^4)$$

and leave it at that.[28]

# 25 Final Comments About Taylor Series

## 25.1 Equality of Functions to their Taylor Series

There is an important theoretical question that we have avoided so far. We have discussed how to calculate Taylor polynomials and Taylor series, and we have discussed how to determine intervals of convergence of these series. However, we have never actually proved that a convergent Taylor series is *equal* to the function it is derived from. In fact, we *cannot* prove this; it isn't necessarily true at all! There is a famous (or infamous?) counterexample: consider the function

$$g(x) = \begin{cases} e^{-1/x^2} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0. \end{cases}$$

Using the definition of the derivative, it is possible to show that $g^{(n)}(0) = 0$, for all $n$. Therefore the Maclaurin series of $g(x)$ is simply the zero function. This, obviously, converges for all $x$, and yet doesn't match $g(x)$ for any $x \neq 0$!

Well, that's disturbing. However, in practice we are *extremely* unlikely to encounter this sort of behaviour. To fully understand why, you'd need to study some complex analysis, but in practice it is quite safe to assume that if a Taylor series of a function $f(x)$ converges, then it converges *to* $f(x)$.

## 25.2 Derivation of the Second-Derivative Test

In Section 7 we stated the Second-Derivative Test, and mentioned that its proof would depend upon concepts in the second half of the course. We now have those concepts, so we can explain why the Test works.

[28] This is not an assumption that we can easily justify. However, the intent is to make clear that we have included all of the terms up to and including 3$^{\text{rd}}$-order, and none of the terms of order higher than 3. For that purpose the notation works well, and the questionable assumption causes no problems.

Suppose we have identified a critical point $P_0 = (x_0, y_0)$ for an infinitely-differentiable function $f(x, y)$, and let $\Delta f = f(x, y) - f(P_0)$. By the definitions of local maxima and minima (in section 7), $f(P_0)$ will be a maximum if $\Delta f < 0$ for all points near $P_0$, and it will be a minimum if $\Delta f > 0$ for all points in some neighbourhood of $P_0$. If, instead, the sign of $\Delta f$ changes for different values of $x$ and $y$ near $P_0$, then the point must be a saddle point.

To determine how $\Delta f$ behaves near $P_0$, we consider the Taylor series of $f$ centered at that point (with the notation $h = x - x_0, \quad k = y - y_0$):

$$f(x, y) = f(P_0) + f_x(P_0)h + f_y(P_0)k + \frac{1}{2}f_{xx}(P_0)h^2 + f_{xy}(P_0)hk + \frac{1}{2}f_{yy}(P_0)k^2 + \dots \quad (48)$$

We can subtract $f(P_0)$ from both sides to get an expression for $\Delta f$. Also, since $P_0$ is a critical point (and we've assumed $f$ to be infinitely differentiable), we know that $f_x(P_0) = 0$ and $f_y(P_0) = 0$, so this reduces to

$$\Delta f = \frac{1}{2}\left[f_{xx}(P_0)h^2 + 2f_{xy}(P_0)hk + f_{yy}(P_0)k^2\right] + \dots \quad (49)$$

Therefore, in order to investigate the sign of $\Delta f$, we can look at the expression in the brackets (if we stay close to the point $P_0$, then the third-order terms should be negligible - as long as the expression in the brackets isn't zero). Actually, we can make the problem a little bit clearer if we make a couple of minor changes: first multiply Equation (49) by $\frac{2}{k^2}$ :

$$\frac{2}{k^2}\Delta f = \left[f_{xx}(P_0)\left(\frac{h}{k}\right)^2 + 2f_{xy}(P_0)\left(\frac{h}{k}\right) + f_{yy}(P_0)\right] + \dots \quad (50)$$

You can now see that the expression in the brackets is a quadratic in $\frac{h}{k}$. If we now relabel the constants as $A = f_{xx}(P_0)$, $B = f_{xy}(P_0)$, $C = f_{yy}(P_0)$, and $\alpha = \frac{h}{k}$, we have

$$\frac{2}{k^2}\Delta f \approx A\alpha^2 + 2B\alpha + C.$$

Why is this useful? It's because we already know a lot about how quadratic functions behave. In particular, we know that if a quadratic has distinct real roots, then its curve crosses the horizontal axis, so the function takes on both positive and negative values. On the other hand, if it has *no* real roots, then the curve never crosses the axis, so the function's values are either

always positive or always negative. This is exactly the information we were seeking about $\Delta f$!

This particular quadratic changes sign for different values of $\alpha$ only if $B^2 - AC > 0$. We can obviously flip that around to say that the quadratic (and, therefore, $\Delta f$) is of *one sign only* if $AC - B^2 > 0$, and that leads directly to the statement of the Test:

**The Second-Derivative Test for Local Extrema**

Suppose $P_0$ is a critical point of a function $f(x, y)$, and suppose that the second-order partial derivatives of $f$ are continuous in some neighbourhood of $P_0$.
Let $D(x, y) = f_{xx}f_{yy} - (f_{xy})^2$.

- If $D(P_0) > 0$, then $f$ has an extremum at $P_0$.
  - If $f_{xx}(P_0) < 0$ then this extremum is a maximum, whereas if $f_{xx}(P_0) > 0$ then it is a minimum.[a]
- If $D(P_0) < 0$, then $f$ does *not* have an extremum at $P_0$ (that is, it has a saddle point instead).
- If $D(P_0) = 0$, the test gives no conclusion.

[a] We could use either $f_{xx}$ or $f_{yy}$ for this part of the test; since at an extremum the concavity will be the same along any cross section.

Why does the test fail if $D(P_0) = 0$? In this case, it may be that all of the second-order derivatives are zero, in which case the second-order approximation is a horizontal plane. At the very least, there are non-zero values of $h$ and $k$ which give a value of zero to the expression $f_{xx}(P_0)h^2 + 2f_{xy}(P_0)hk + f_{yy}(P_0)k^2$, and so we would have to consider the third-order terms in the Taylor expansion of $f$ in order to determine the sign of $\Delta f$ (or, more practically, find another strategy altogether).

## Part III

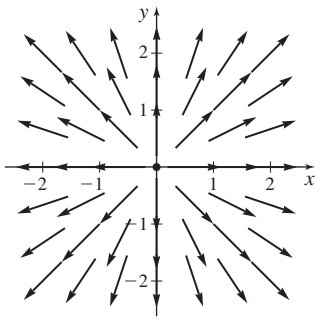
# A Brief Introduction to the Calculus of Vector Fields

*Vector fields* are mappings from vectors to vectors: $\vec{F} : \mathbb{R}^2 \to \mathbb{R}^2$, or more generally, $\vec{F} : \mathbb{R}^n \to \mathbb{R}^n$.

We'll concentrate on the 2D variety in this discussion. Since they have two inputs and two outputs, they can be described as ordered pairs of scalar fields: $\vec{F}(x, y) = (F_1(x, y), F_2(x, y))$. We can plot a vector field by simply sketching a sample set of vectors.

**Example 1:** Let $\vec{F}(x, y) = \left( \frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}} \right)$. This looks complicated for a first example, but at every point $(x, y)$ we can see that $\vec{F}(x, y)$ is a unit vector directed away from the origin (since $\left|\left|\vec{F}\right|\right| = \frac{1}{\sqrt{x^2+y^2}} \sqrt{x^2 + y^2} = 1$). Plotting a few of these, we obtain this representation:

Figure 43:




Here we might call the origin a *source*. If the vectors were reversed we would instead call it a *sink*.

**Example 2:** Let $\vec{G}(x, y) = (x, -y)$. Pick a few points, and sketch the corresponding vectors:

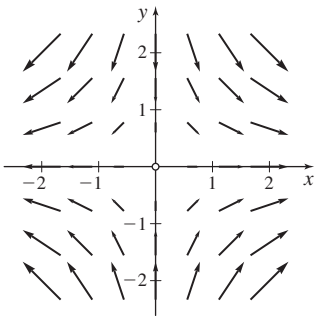



Figure 44:

**Example 3:** Let $\vec{H}(x, y) = (y, 2xy)$. We can try the same strategy, but it's harder to make sense of what we see:

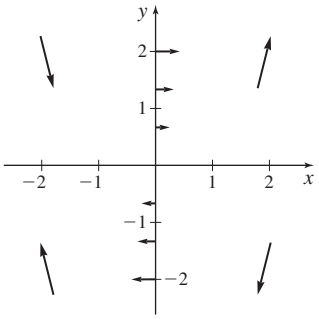



Figure 45:

Maple's "fieldplot" command gives us a cleaner look (below), but plotting by hand does at least force us to notice that there is a line of zero vectors along the $x$-axis.

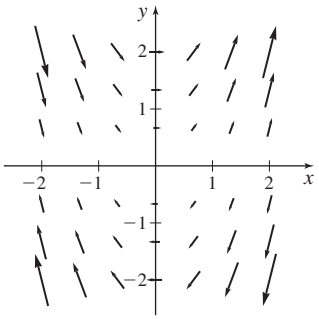
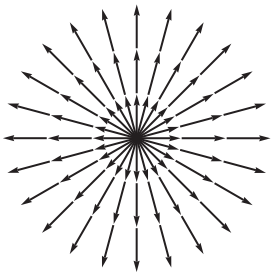



Figure 46:

We do have an alternative way of graphing vector fields, but it requires calculations which can't always be performed by hand. A field line of $\vec{F}(x, y)$ is a curve whose tangent coincides with $\vec{F}$ at each point. We can sometimes guess at what these should look like from the direction fields; for example for $\vec{F}(x, y)$ in Example 1 they are straight lines:

Figure 47: 

To actually calculate the equations of the field lines, we give them parametric representations. If $\vec{r}(t) = (x(t), y(t))$ is a field line of $\vec{F}(x,y) = (F_1, F_2)$, then we want to have $\vec{r}\,'(t) = \vec{F}$, so we set $x' = F_1$ and $y' = F_2$. This gives us a pair of coupled differential equations, which we then attempt to solve to find $x(t)$ and $y(t)$.

**Example 4:** For $\vec{F}(x,y)$ of Example 1, we set

$$\frac{dx}{dt} = \frac{x}{\sqrt{x^2+y^2}}, \qquad \frac{dy}{dt} = \frac{y}{\sqrt{x^2+y^2}}.$$

We can obtain a differential equation for $y$ as a function of $x$ by eliminating $dt$:

$$dt = \frac{\sqrt{x^2+y^2}}{x}dx = \frac{\sqrt{x^2+y^2}}{y}dy$$

$$\Longrightarrow \qquad \frac{dx}{x} = \frac{dy}{y} \qquad (\text{for } (x,y) \neq (0,0)$$

$$\Longrightarrow \qquad \frac{dy}{dx} = \frac{y}{x}.$$

We can solve this DE by separating the variables (which in this case actually means going back one step) and integrating:

$$\int \frac{dx}{x} = \int \frac{dy}{y}$$

$$\Longrightarrow \ln|x| = \ln|y| + C_1$$

$$\Longrightarrow |x| = e^{\ln|y|+C_1} = e^{C_1}|y|$$

$$\Longrightarrow x = C_2 y \qquad \left(\text{where } C_2 = \pm e^{C_1}\right)$$

$$y = kx \qquad (\text{where } k \text{ is a constant}).$$

These are straight lines through the origin, as expected!

**Example 5:** Repeating this for $\vec{G}(x, y) = (x, -y)$, we find hyperbolae: $y = k/x$, so the field lines look like this:

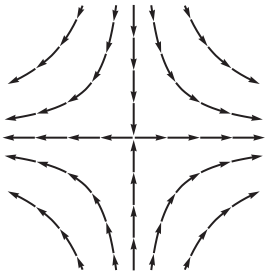


Figure 48:

**Example 6:** For $\vec{H}(x, y) = (y, 2xy)$, we get parabolas: $y = x^2 + k$. The work we did in Example 3 tells us the direction of motion; in fact we have a line of sinks and sources along the $x$-axis (sinks for $x < 0$, sources for $x > 0$, with the origin being both at once).

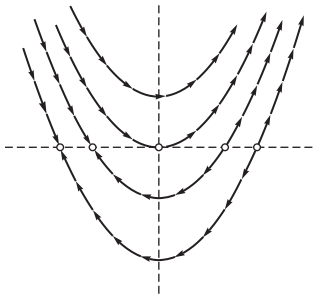


Figure 49:

## Divergence & Curl

Can we differentiate a vector field? Well, consider a 2-variable one: $\vec{F} = (F_1(x, y), F_2(x, y))$. There are four partial derivatives, and it turns out that there are two combinations of them which are of particular interest:

- The sum of the "straight" derivatives, $\dfrac{\partial F_1}{\partial x} + \dfrac{\partial F_2}{\partial y}$, is called the divergence of $\vec{F}$, sometimes written as "Div $\vec{F}$".
- The difference of the "crossed" derivatives, $\dfrac{\partial F_2}{\partial x} - \dfrac{\partial F_1}{\partial y}$, is called the curl of $\vec{F}$, sometimes written as "Curl $\vec{F}$".

Div $\vec{F}$ can be interpreted as a measure of local expansion, while Curl $\vec{F}$ can be interpreted as a measure of local rotation. It's helpful to think of $\vec{F}$ as representing velocities of a fluid.

Since we're working in 2D, let's suppose it describes velocities on the surface of a body of water. Imagine placing a drop of dye in the water. If Div $\vec{F} > 0$, then the drop will start to expand. If Curl $\vec{F} > 0$, then it will start to rotate counterclockwise[29] (so we might have to use a small X instead of a simple droplet to see this).

**Example 7:** For our fields $\vec{F}$ and $\vec{G}$ from earlier, we find that both the divergence and the curl are zero (for both fields). We call such fields *incompressible* and *irrotational*. For our third example, $\vec{H} = (y, 2xy)$, though, we find that

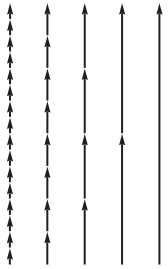
$$\text{Div } \vec{H} = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(2xy) = 2x,$$

so we have expansion in the right half of the plane and contraction on the left. Meanwhile,

$$\text{Curl } \vec{H} = \frac{\partial}{\partial x}(2xy) - \frac{\partial}{\partial y}(y) = 2y - 1,$$

so we have counterclockwise rotation when $y > 1/2$ and clockwise rotation when $y < 1/2$.

**Example 8 (one last simple one):** Let $\vec{F}(x, y) = (0, x)$, for $x \geq 0$. This might represent the surface currents in a river, near the left riverbank:

Figure 50: 

Here we have Div $\vec{F} = 0$, while Curl $\vec{F} = 1$. Objects dropped in the water would travel straight downstream, rotating counterclockwise along the way!

[29]The choice of $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$ instead of $\frac{\partial F_1}{\partial y} - \frac{\partial F_2}{\partial x}$ may look arbitrary, but the sign of this difference tells us the direction of rotation. Counterclockwise rotation occurs when $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} > 0$, and so the curl is defined so that positive curl corresponds to rotation in the "positive" direction.

**Extension to Three Dimensional Fields**

In $\mathbb{R}^3$, a vector field will comprise three functions of three variables:

$$\vec{F} = \left(F_1\left(x,y,z\right), F_2\left(x,y,z\right), F_3\left(x,y,z\right)\right).$$

You can probably guess at the formula for the divergence; it's simply $\dfrac{\partial F_1}{\partial x} + \dfrac{\partial F_2}{\partial y} + \dfrac{\partial F_3}{\partial z}$. The curl, though, is a bit more complicated. To describe rotation in three dimensions, we need to specify both a magnitude and an orientation, so the curl needs to be defined as a *vector*. Here it is:

$$\text{Curl } \vec{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right).$$

The plane of rotation is orthogonal to the Curl vector, with the direction of rotation given by the right-hand rule:

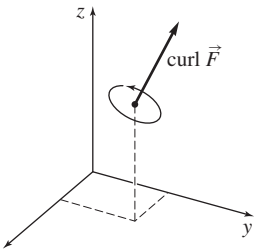



Figure 51:

The formula looks overwhelming, but there is a clear pattern. To find the $x$-component of the Curl, we need to differentiate the $z$-component of $\vec{F}$ with respect to $y$, and differentiate the $y$-component of $\vec{F}$ with respect to $z$ (and subtract them). A different notation might make it clearer; perhaps we could write $\text{Curl}_x = Z_y - Y_z$. The other components are $\text{Curl}_y = X_z - Z_x$ and $\text{Curl}_z = Y_x - X_y$. Another way to see the pattern is to consider that the three-dimensional curl should reduce to the two-dimensional version in certain circumstances. For example, a 2D vector field $\vec{F}\left(x,y\right) = \left(F_1\left(x,y\right), F_2\left(x,y\right)\right)$ can be viewed as a 3D field with a zero $z$-component: $\vec{F}\left(x,y,z\right) = \left(F_1\left(x,y\right), F_2\left(x,y\right), 0\right)$. Our 3D definition of Curl gives

$$\text{Curl } \vec{F} = \left(0, 0, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right),$$

and this confirms that the plane of rotation is the $xy$-plane, and the direction is counterclock-

wise when $\dfrac{\partial F_2}{\partial x} - \dfrac{\partial F_1}{\partial y} > 0$ (by the right-hand rule).

**Nabla Notation**

It is possible to express the divergence and curl in an extremely concise notation. We first define the operator $\nabla$ to be $\nabla = \left(\dfrac{\partial}{\partial x}, \dfrac{\partial}{\partial y}\right)$. The idea is to treat this as a vector to begin with, and apply the derivatives as soon as we have something to apply them to. For example, we can now interpret the gradient vector this way:

$$\nabla f = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) f$$

$$= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right).$$

With this idea, we can write

$$\text{Div } \vec{F} = \nabla \cdot \vec{F}$$

and, with a bit of creativity,

$$\text{Curl } \vec{F} = \nabla \times \vec{F}$$

We'll leave it to you to confirm that these give the appropriate formulas (note that $\nabla \times \vec{F}$ can be expanded as $\begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix}$).

## Line Integrals (a.k.a. Path Integrals)

Can we integrate vector fields? You might want to think about what that question actually means; what quantities would you wish to add up? There is, in fact, one important kind of vector field integral, but the idea probably won't be obvious at this point. To introduce the concept, let's go back to scalar fields once more; we've discussed how to integrate two-variable functions over two-dimensional regions, but it is also possible to integrate them *along curves*. This is related to the idea of constrained optimization - we discussed how to find the extreme values of a function along a specific curve, and now we'll discuss how to sum *all* of its values along the curve. From a geometric viewpoint, the goal is to find the area of what one might call a "curtain" lying above a curve $C$ with height defined by $f(x, y)$.

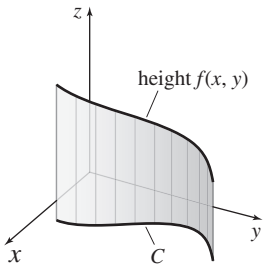


Figure 52:

We'll need to express the curve $C$ parametrically: we'll use $\vec{r}(t) = (x(t), y(t))$, $t \in [a, b]$. We subdivide $C$ into segments of length $ds$ (this might seem familiar; we did the same thing to develop the arc length formula in Math 117). This can be imagined as dividing our curtain into short panels, of area

$$dA = f(x, y)\, ds.$$

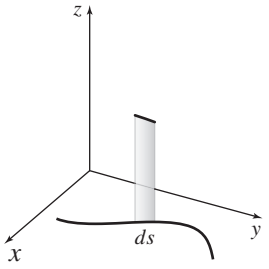


Figure 53:

Summing these, we'll find

$$\text{Area } = \int f\, ds.$$

Now what's $ds$? It's the same as in our arc length formula!

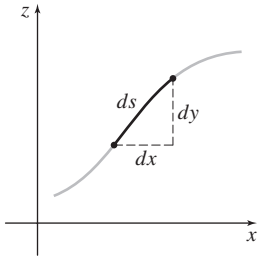


Figure 54:

$$ds = \sqrt{dx^2 + dy^2}$$

Factoring out $dt$, we obtain

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt,$$

which we may write as

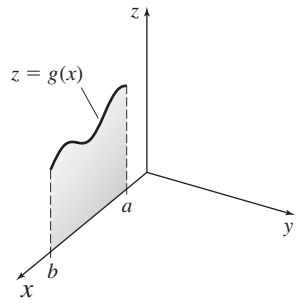
$$ds = \left|\left|\vec{r}\,'(t)\right|\right| dt.$$

We therefore define the *line integral of* $f(x,y)$ *along the curve* $C$ as follows:

$$\boxed{\int_C f\, ds = \int_a^b f\left(\vec{r}(t)\right) \left|\left|\vec{r}\,'(t)\right|\right| dt}$$

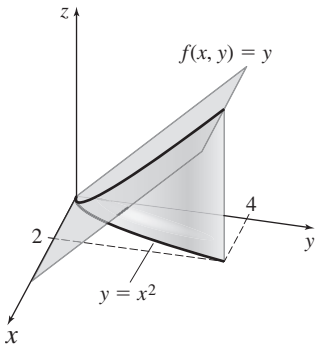
Note that in the special case where $f(x,y) = g(x)$ and our curve is the $x$-axis (parameterized as $\vec{r}(t) = (t,0)$), we get an ordinary integral:

$$\begin{aligned} \text{area} &= \int_C f\, ds \\ &= \int_a^b g(t)\, dt \\ &= \int_a^b g(x)\, dx \end{aligned}$$




Also notice that as usual, our definition says nothing about areas, and so we can interpret a line integral more generally as a kind of sum of all of the values of $f$ which occur along the curve $C$. In particular, if the function is simply $f(x,y) = 1$, then we get the length of the curve, so the arc-length formula can now be seen as a special case of the path integral!

**Example 9:** Evaluate $\int_C f\,ds$, where $f(x,y) = y$ (this is a plane inclined at 45°) and $C$ is the part of the parabola $y = x^2$ between $x = 0$ and $x = 2$.




Parameterize $C$: $\vec{r}(t) = \left(t, t^2\right), \qquad t \in [0,2]$.

Then $\vec{r}\,'(t) = (1, 2t)$, and so $||\vec{r}\,'(t)|| = \sqrt{1+4t^2}$.

Also, along the curve $C$, $f(x,y) = y = t^2$.

$$\Longrightarrow \int_C f\,ds = \int_0^2 t^2\sqrt{1+4t^2}dt$$

$$\approx 8.47 \quad \text{(using a calculator)}$$

**Line Integrals of Vector Fields**

Integration of vector fields is motivated by the following question:

Suppose a vector field $\vec{F}$ is understood to represent a force (it could be an electric field or a gravitational field, or any other kind of force field). Suppose also that a particle moves along a path $C$, experiencing the force $\vec{F}$ as it moves. How much *work* does the force field perform on the particle?

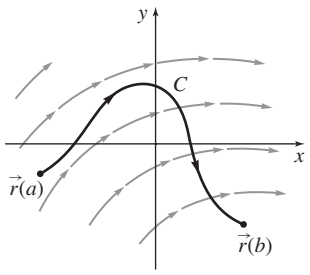



Figure 55:

We know from elementary physics that for motion in a straight line, work = force×displacement. If the motion is not in a straight line, we know that we need vectors, and in that case $W = \vec{F}\cdot\vec{d}$, but this still requires that $\vec{F}$ be constant! So, in the usual fashion, we break the problem into small pieces.

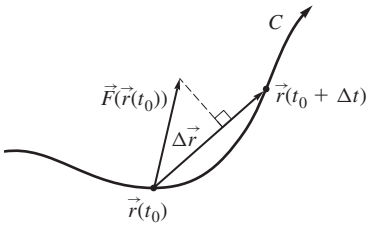



Figure 56:

As our particle moves from $\vec{r}\,(t_0)$ to $\vec{r}\,(t_0 + \Delta t)$, resulting in a small displacement $\Delta\vec{r}$, it is subjected to a nearly constant force $\vec{F}\,(\vec{r}\,(t_0))$. The work performed by $\vec{F}$ on the particle over this part of the path is

$$\Delta W \approx \vec{F} \cdot \Delta\vec{r}.$$

What's $\Delta\vec{r}$? Well, $\dfrac{d\vec{r}}{dt} = \lim\limits_{\Delta t \to 0} \dfrac{\Delta\vec{r}}{\Delta t}$, so we may write $\Delta\vec{r} \approx \dfrac{d\vec{r}}{dt}\Delta t$. Hence $\Delta W \approx \vec{F} \cdot \Delta\vec{r} \approx \vec{F} \cdot \dfrac{d\vec{r}}{dt}\Delta t$. Now we just add all these together, and let $\Delta t \to 0$:

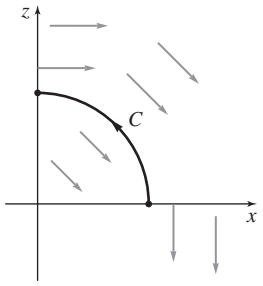
$$\text{Work} = \int_a^b \vec{F} \cdot \frac{d\vec{r}}{dt} dt.$$

We call this the line integral of $\vec{F}$ along $C$, and we have a shorthand notation for it (we just cancel the differentials $dt$, and remove the references to $t = a$ and $t = b$ from the limits of integration):

$$\int_C \vec{F} \cdot d\vec{r} := \int_a^b \vec{F} \cdot \frac{d\vec{r}}{dt} dt = \int_a^b \vec{F}\,(\vec{r}\,(t)) \cdot \vec{r}\,'\,(t)\, dt$$

(note that after we perform the dot product in this expression, we're left with an ordinary single integral).

**Example 10:** Evaluate $\int_C \vec{F} \cdot d\vec{r}$, where $\vec{F}(x, y) = (y, -x)$ and $C$ follows the unit circle from $(1, 0)$ to $(0, 1)$.




**Solution:** $C$ can be parameterized as $\vec{r}(t) = (\cos t, \sin t)$, with $t \in [0, \pi/2]$.

With this, $\vec{F}(\vec{r}(t)) = (\sin t, -\cos t)$, and $\vec{r}\,'(t) = (-\sin t, \cos t)$. Therefore

$$\int_C \vec{F} \cdot d\vec{r} = \int_0^{\pi/2} (\sin t, -\cos t) \cdot (-\sin t, \cos t)\, dt$$

$$= \int_0^{\pi/2} \left(-\sin^2 t - \cos^2 t\right) dt$$

$$= \int_0^{\pi/2} (-1)\, dt = -\frac{\pi}{2}$$

(the work is negative because the motion is *opposed* by the force in this example).

**Example 11:** What if we move from $(1, 0)$ to $(0, 1)$ along a straight line instead (using the same $\vec{F}$)?

Then we can use

$$\vec{r}(t) = (1, 0) + t(-1, 1)$$

$$= (1 - t, t), \qquad t \in [0, 1].$$

Now $\vec{F}(\vec{r}(t)) = (y, -x) = (t, t - 1)$, and $\vec{r}\,'(t) = (-1, 1)$. Therefore

$$\int_C \vec{F} \cdot d\vec{r} = \int_0^1 (t, t - 1) \cdot (-1, 1)\, dt = \int_0^1 (-t + t - 1)\, dt = -1.$$

Notice that even though we used the same vector field $\vec{F}$, moving from $(1, 0)$ to $(0, 1)$ along different paths resulted in different amounts of work. In applications, though, we often encounter fields for which the value of a line integral between any two fixed points is independent of the path taken. Such fields are called *conservative*, because they can be interpreted as corresponding to forces which obey the Law of Conservation of Energy.

A couple of final remarks on notation:

- If $C$ happens to be a closed path (a loop), we may use the special notation $\oint_C \vec{F} \cdot d\vec{r}$.

- Observe that

$$\int_C \vec{F} \cdot d\vec{r} = \int_C \vec{F} \cdot \vec{r}\,'(t)\,dt$$

$$= \int_C (F_1, F_2)\left(\frac{dx}{dt}, \frac{dy}{dt}\right) dt$$

$$= \int_C \left(F_1 \frac{dx}{dt} dt + F_2 \frac{dy}{dt} dt\right)$$

$$= \int_C F_1 dx + F_2 dy$$

The notation found in the final line here is used heavily in most textbooks. It's a convenient notation to start with if our path is a zig-zag, made up of segments parallel to the $x$- and $y$-axes (and if the vector field $\vec{F}$ is conservative then we can *choose* the path to be of this type). However, if the vector field is *not* conservative, and our path is *not* made up of horizontal and vertical segments, then we will need to parameterize the curve, and so we'll need to go back to the form $\displaystyle\int_a^b \vec{F}\left(\vec{r}(t)\right) \cdot \vec{r}\,'(t)\,dt$. If in doubt, just remember that $\displaystyle\int_C F_1 dx + F_2 dy$ *means* $\displaystyle\int_C \vec{F} \cdot d\vec{r}$.

---

Those are the basic concepts of vector calculus. There are more (the next we would discuss if we had time would be *flux* and *surface integrals*). There are also some very important theorems relating some of these concepts (Green's Theorem, Gauss' Theorem, and Stokes' Theorem), but we'll have to leave these for another course.