

ESS 314

1. Mathematical and Physical Tools by John R. Booker and Gerard Roe

Physics is the set of rules by which we think the world works; mathematics is the language we use to compactly and precisely describe these rules and use the rules to make predictions. Our predictions may be of value themselves or be a test of whether our physics is correct. This is an overview of tools used in this course. Most should be familiar from course pre-requisites. Some are extensions of this assumed knowledge.

1.1 Mathematical tools

Numbers: real and imaginary

Our most basic need is to describe how big something is. How long is it; how much does it weigh; how fast is it moving? This is what positive numbers are for. Positive numbers can be added or subtracted. Subtleties that are not very important to us is that some numbers are integers (whole numbers like 1, 13 or 65), others are rational (the ratio of two integers: 1.12 is the ratio of 112 and 100) and some are irrational (cannot be specified as the ratio of two integers, but may be the ratio of physical quantities described to infinite accuracy. Examples of irrational numbers are $\sqrt{2}$ or π (pi), the ratio of the circumference to the diameter of a circle. If a number requires an infinite number of places after the decimal point to express it completely, then it is irrational.

Subtraction can also be thought of as the addition of a positive and a negative number. Negative numbers are useful in representing physical quantities like the charge of an electron. Negative numbers introduce an important complication because their square root is not a positive or negative number. The most fundamental square root of a negative number is the square root of -1 , which is commonly called “i” or (in engineering, “j”). Square roots of negative numbers are said to be “imaginary”, while ordinary numbers are said to be “real”. We shall see that imaginary numbers have a very real meaning in the physical world.

Quantities that are the sum of a real and an imaginary number are said to be “complex” numbers. If complex numbers are equal, their real and imaginary parts are separately equal. Thus $a+ib = c+id$ implies that $a = c$ and $b = d$. The notation $\text{Re}\{c\}$ and $\text{Im}\{c\}$ means the real and imaginary parts of a complex number c . For $c = a+ib$, $\text{Re}\{c\} = a$ and $\text{Im}\{c\} = b$. The “complex conjugate” of a complex number involves reversing the sign of its imaginary part. The complex conjugate of $a+ib$ is written $(a+ib)^* = a-ib$.

Multiplication of real integers is first learned in grade school. It gets slightly more complicated when we consider the multiplication of rational, negative, imaginary and complex numbers, but the concepts are all straightforward generalizations. The following in which a, b, c and d are positive real numbers summarizes the various rules

$$\begin{aligned}(a)(b) &= (b)(a) = ab \\(a)(-b) &= (-a)(b) = -ab \\(-a)(-b) &= ab \\(a)(ib) &= (ia)(b) = i(ab) \\(ia)(ib) &= -ab \\(a+ib)(c+id) &= (ac-bd) + i(ad+bd)\end{aligned}$$

For real or imaginary numbers, the order in a multiplication is unimportant. This is not true for complex numbers. This is easily seen by

$$\begin{aligned}(a+ib)(c+id)^* &= (a+ib)(c-id) = (ac+bd) - i(ad-bc) \\&= [(ac+bd) + i(ad-bc)]^* = [(c-id)(a+ib)]^* = [(c+id)^*(a+ib)]^*\end{aligned}$$

Thus reversing the order of multiplication of complex numbers results in the complex conjugate of the original result.

A special case of complex multiplication that comes up very often is

$$(a+ib)(a+ib)^* = (a+ib)(a-ib) = a^2 + b^2 = |(a+ib)|^2$$

where the $|\quad|$ symbol means “magnitude (or size) of”. Thus a complex number times its conjugate is the square of its magnitude.

Vectors

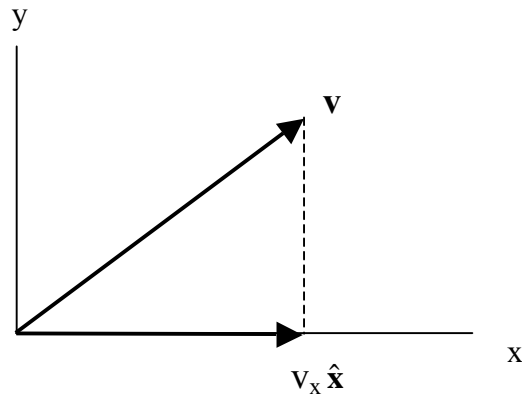
A number is sufficient when only size is important, but numbers are cumbersome when used to describe things that have direction as well as magnitude. Vectors were invented to simplify the situation. A vector is an arrow with both length (magnitude) and direction. For instance, the velocity of an object is a vector that points in the direction that it is moving and has a length equal to its speed (the rate of change of distance with time). In this text, we will write vectors with bold face symbols. The symbols **x** or **r** will usually mean position and **v** or **u** will almost always mean velocity. In lecture, we will use symbols like \vec{v} or \underline{v} because bold face is hard to write on a black or white board.

A vector can be written in terms of its components along orthogonal “coordinate axes”

$$\mathbf{v} = v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}} + v_z \hat{\mathbf{z}}$$

In this case, $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are “unit vectors” that point along the x, y and z axes of a rectangular coordinate system and have a length equal to 1 in whatever units the length or

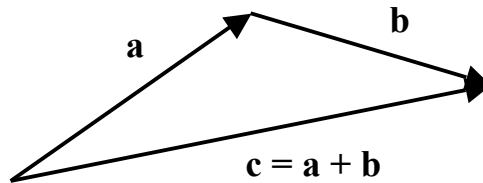
magnitude of vector \mathbf{v} is measured. The “components” v_x , v_y and v_z are the projections of the vector on each of the three axes. Graphically, v_x is related to \mathbf{v} this way:



The square of the length or magnitude of \mathbf{v} in terms of its components is

$$|\mathbf{v}|^2 = v_x^2 + v_y^2 + v_z^2$$

Vectors can be added as shown in this illustration:

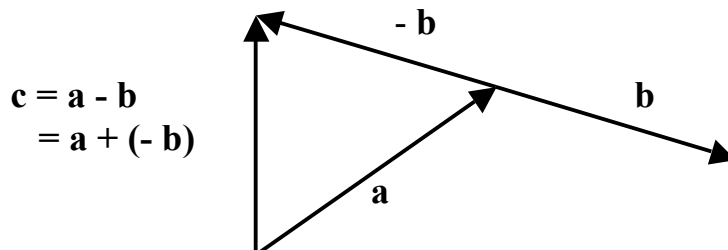


Another way to get to the same result is to realize that the components of the sum vector are the sum of the components of the two vectors:

$$\mathbf{a} + \mathbf{b} = (a_x + b_x) \hat{\mathbf{x}} + (a_y + b_y) \hat{\mathbf{y}} + (a_z + b_z) \hat{\mathbf{z}}$$

The graphical way is more powerful because it does not rely on any imposed coordinate system. Often, there is no reason to impose a particular coordinate system on a physical situation.

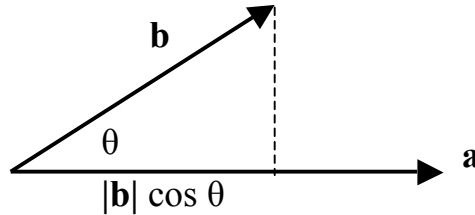
Vector subtraction is easily understood as the sum of a positive and a negative vector. The negative vector has the same magnitude as the positive vector, but points in the opposite direction. For the same vectors \mathbf{a} and \mathbf{b} used above, we have:



Dot product of vectors

Vectors can also be multiplied. There are two ways to do this. One, called the “dot” or “scalar” product results in a number (sometimes called a “scalar” to distinguish it from a vector); the second called the “cross” or “vector” product results in a vector.

If the angle between vectors **a** and **b** is θ , the dot product **a**·**b** involves projecting **b** onto the direction of **a** like this



and then multiplying the length of this projected vector times the length of **a**. Thus

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos \theta$$

The dot product is maximum when **a** and **b** are parallel and zero when they are perpendicular. Unit vectors of a rectangular coordinate system have

$$\begin{aligned}\hat{\mathbf{x}} \cdot \hat{\mathbf{x}} &= \hat{\mathbf{y}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} = 1 \\ \hat{\mathbf{x}} \cdot \hat{\mathbf{y}} &= \hat{\mathbf{x}} \cdot \hat{\mathbf{z}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{z}} = 0\end{aligned}$$

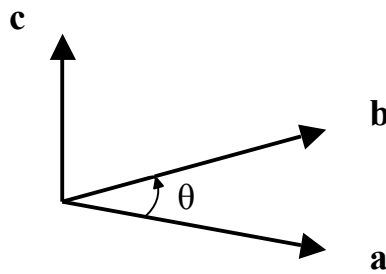
These relations allow one to express the dot product in terms of vector components as

$$\mathbf{a} \cdot \mathbf{b} = (a_x \hat{\mathbf{x}} + a_y \hat{\mathbf{y}} + a_z \hat{\mathbf{z}}) \cdot (b_x \hat{\mathbf{x}} + b_y \hat{\mathbf{y}} + b_z \hat{\mathbf{z}}) = a_x b_x + a_y b_y + a_z b_z$$

It should be obvious that the roles of **a** and **b** can be interchanged in computing the dot product. In fact, the dot product depends only on the angle between two vectors and not on their absolute directions. Note the dot product of a vector with itself is just the square of its magnitude.

Cross product of vectors

The geometry of the cross product **c** = **a** × **b** looks like this:



Vector **c** is perpendicular to both **a** and **b** and θ is the angle between vectors **a** and **b**. By convention, the direction of **c** is given by the “right hand rule”: Hold your right hand with the thumb pointing up and your index finger in the direction of **a**. If you can rotate your index finger counter-clockwise less than half a turn to the direction of **b**, then your thumb points in the direction of **c**. If you have to rotate clockwise to get to **b**, then **c** is opposite to the direction of your thumb. This implies that the roles of **a** and **b** cannot be arbitrarily interchanged. In fact

$$\mathbf{b} \times \mathbf{a} = -\mathbf{a} \times \mathbf{b}$$

The magnitude of a cross product is given by

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin \theta$$

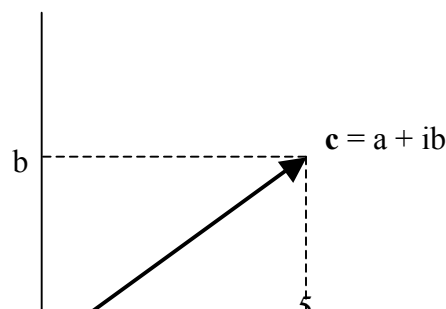
Thus the cross product is maximum when **a** and **b** are perpendicular to each other and zero when they are parallel.

Expressing **c** in terms of the components of **a** and **b** is an algebraic mess and you need to be careful about the “handedness” of your coordinate system and whether it is “Cartesian” (standard orthogonal) or “curvilinear” (such as spherical or cylindrical). The convention we use in this course is that, unless specifically stated otherwise, coordinates are always right-handed and Cartesian. If the index finger of your right hand points along the x axis and your middle finger is bent at a right angle to point along the y axis, the z axis will be in the direction of your thumb. You can use the properties of the 3×3 determinant (if you are familiar with them) to write

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = \hat{\mathbf{x}}(a_y b_z - a_z b_y) - \hat{\mathbf{y}}(a_x b_z - a_z b_x) + \hat{\mathbf{z}}(a_x b_y - a_y b_x)$$

This may seem un-necessarily complex, but the cross product turns out to be extremely useful in the context of the physics of planetary atmospheres and oceans because the subtle, but critically important effect of planetary rotation on winds and currents is easily expressed using the cross product in a way that is independent of coordinate system. The forces on a charged particle influenced by a magnetic field in Earth’s space environment are also much easier to work with using the cross product.

Finally, complex numbers are usefully represented as two-dimensional vectors with the x imaginary axis

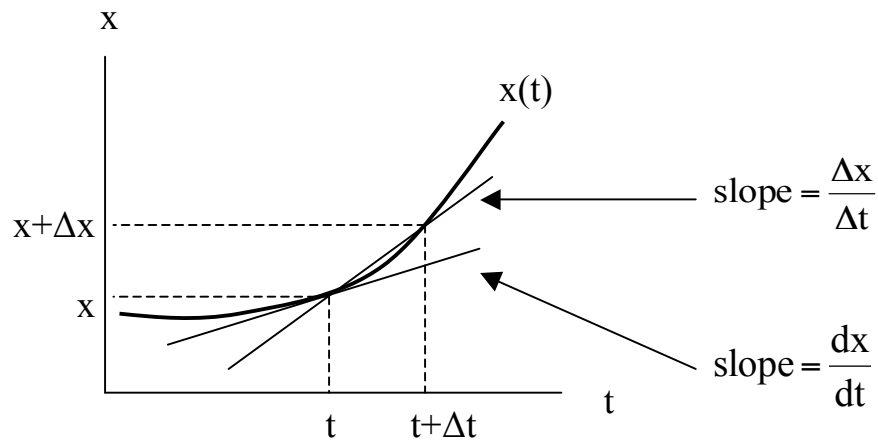


and y axes replaced by the real and imaginary axes. This is called the “complex plane”.

Looking back at the definitions for the magnitude of a complex number and a vector, you see that they are identical when a is the real “component” and b is the imaginary “component”. This called the “rectangular” representation of a complex number.

Functions and derivatives

Physical properties, such as mass and charge can change with position and time. When this happens, we say that they are “functions” of space and time and we need tools for dealing with their rates of change. Suppose that position x depends only on time t in such a way that there is only one possible value of x for each t . $x(t)$ is thus a single-valued function of its argument t .



If we calculate x at t , at a later time $t+\Delta t$, x will have changed to $x(t+\Delta t) = x(t) + \Delta x$. The average rate of change of x with respect to t (the speed) over the time interval Δt is

$$\frac{x(t + \Delta t) - x(t)}{(t + \Delta t) - t} = \frac{x(t) + \Delta x - x(t)}{\Delta t} = \frac{\Delta x}{\Delta t}$$

which is the slope of a line through the points (x, t) and $(x+\Delta x, t+\Delta t)$. Repeating this calculation for successively smaller values of Δt , this slope approaches the slope of a line tangent to the $x(t)$ curve at the point $x(t)$. The quantity

$$\frac{dx}{dt} = \text{limit as } \Delta x \rightarrow 0 \text{ of } \left(\frac{\Delta x}{\Delta t} \right)$$

is called the derivative of x with respect to t . The derivative of a curve is the slope of the tangent line. The tremendous breakthrough of the late 18th Century giants Newton and Descartes was to figure out how to calculate derivatives of functions without going through this limiting process. This allowed them to write down functional relations between, for instance, position and velocity and was essential to the development of

physics. Ironically, due the speed and large memory of modern computers, derivatives are now often routinely calculated with the old, discrete method.

Higher degree derivatives are simply repeated application of the derivative. For instance

$$\frac{d^2f}{dx^2} = \frac{d}{dx} \left(\frac{df}{dx} \right) \quad \frac{d^3f}{dx^3} = \frac{d}{dx} \left\{ \frac{d^2f}{dx^2} \right\} = \frac{d}{dx} \left\{ \frac{d}{dx} \left(\frac{df}{dx} \right) \right\}$$

Enormous tomes exist describing the many mathematical functions useful in physics. However, most of the equations of physics have solutions that can be closely approximated by a very small number of functions that you really need to know about.

The most important class of functions are exponentials e^t , where e is the irrational number 2.71828... If the argument t is positive, this function grows without bound; if it is negative, it decays to zero; if it is imaginary it oscillates. The exponential function has the remarkable property that its derivative is itself. Thus

$$\frac{d}{dt} e^t = e^t$$

The second class of functions are polynomials, which are sums of terms of the form t^n , where n is a positive or negative integer. Rarely is $|n|$ greater than 2. Polynomials include straight lines, parabolas, hyperbolas, ellipses, etc. The derivative of the n th term of a polynomial is

$$\frac{d}{dt} t^n = nt^{n-1}$$

except for the case $n = 0$ for which

$$\frac{d}{dt} t^0 = \frac{d}{dt} (1) = \frac{d}{dt} (\text{constant}) = 0$$

Since most functions consist of terms combined in various we ways, we need the following rules for calculating the derivatives of combinations:

$$\frac{d}{dx} [f(x) + g(x)] = \frac{df}{dx} + \frac{dg}{dx}$$

$$\frac{d}{dx} [f(x)g(x)] = f \frac{dg}{dx} + g \frac{df}{dx}$$

$$\frac{d}{dx} f(g(x)) = \frac{df}{dg} \frac{dg}{dx}$$

The first rule implies that the derivative of a polynomial with more than one term is the sum of the derivatives of its individual terms. The third “chain rule” is used very often. For instance if we let $g = -\alpha t$ we have

$$\frac{d}{dt} e^{-\alpha t} = \left(\frac{d}{dg} e^g \right) \frac{d}{dt} (-\alpha t) = e^g (-\alpha) = -\alpha e^{-\alpha t}$$

Integrals

Integration is the inverse of differentiation. Differentiation is computation of the slope of a curve; integration is the construction of a curve given its slope. The “indefinite” integral is defined by

$$\int \frac{df(x)}{dx} dx = f(x) + \text{a constant}$$

The indefinite integral is not unique because a constant has a derivative of 0 and so any size constant can be added to the right side of the above relation.

We obviously have

$$\int e^t dt = e^t + \text{a constant}$$

and one can easily show that

$$\int e^{\alpha t} dt = \frac{1}{\alpha} e^{\alpha t} + \text{a constant}$$

Except for the special case with $n = -1$, the indefinite integral of the n th term of a polynomial is

$$\int x^n dx = \frac{x^{n+1}}{n+1} + \text{a constant}$$

as you can easily see by noting that the derivative of the right hand side is x^n .

The special case with $n = -1$ is

$$\int \frac{dx}{x} = \ln(x) + \text{a constant}$$

where the “natural logarithm” $\ln(x)$ (sometimes called the “napierian logarithm” after its discoverer Napier; also sometimes called the base e logarithm) is defined by

$$x = e^{\ln(x)}$$

The natural logarithm is the third type of function you need to know about. Its derivative from the next-to-last equation above is obviously

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

You may be more familiar with the base 10 logarithm defined by

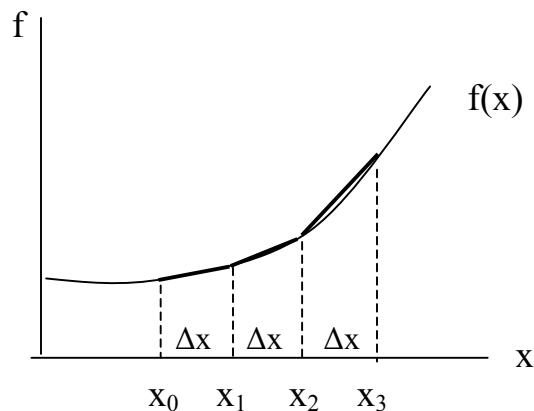
$$x = 10^{\log_{10}(x)}$$

In order to avoid confusion between these two types of logarithms, we always write $\ln(x)$ for the natural logarithm although computer languages such as Matlab, C and Fortran write $\log(x)$. Base 10 logarithms will always be written $\log_{10}(x)$ while computer languages commonly use $\log_{10}(x)$. A useful property of all logarithms easily proved from their definitions is

$$\log(xy) = \log(x) + \log(y)$$

where we have not specified the base because this holds for all bases. Only 35 years ago, base 10 logarithms were extremely important in courses like this because hand-held calculators cost as much as high-end laptop PCs today. Scientists routinely did accurate multiplication using slide rules that mechanically added base 10 logarithms. High school students were taught to do the same thing using base 10 logarithm tables.

You were probably first introduced to integrals as the “area under a curve”. We can approximate $f(x)$ in the figure below with three straight lines.



The area between the function curve and the portion of the x axis between x_0 and x_3 can then be approximated by summing the area of the trapezoids of width Δx . The result is

$$\text{Area} = \left(\frac{f(x_0)}{2} + f(x_1) + f(x_2) + \frac{f(x_3)}{2} \right) \Delta x$$

If we repeat this process with a smaller Δx and more trapezoids, the area will be a better approximation to the area under the curve. The limit as Δx becomes very small and the number of trapezoids becomes very large is the true area under the curve and is called the “definite” integral of $f(x)$:

$$\int_{x_{\min}}^{x_{\max}} \frac{df(x)}{dx} dx = f(x_{\max}) - f(x_{\min})$$

Solution of a simple Ordinary Differential Equation (ODE)

From our point-of-view, the most important use of integration is for the solution of differential equations. An ODE is an equation with at least one term that is a derivative with respect to one variable. A simple, but very important example is

$$\frac{d\rho}{dz} = -\frac{\rho}{H}$$

where ρ is a function only of z and H is a constant. ODE’s of this form come up over and over again. This ODE involves only a first derivative and is said to be of “1st degree”. Its solution is actually obvious, because the equation states that the derivative of the function ρ is equal to itself (scaled by the constant H). We have already noted that the function whose derivative is equal to itself is the exponential. So the solution must be an exponential. However, it is instructive to solve this equation this time in a more formal way using integration.

Remembering that the derivative is really the ratio of a small change in ρ due to a small change in z , we can multiply the above equation by dz and divide it by ρ to obtain

$$\frac{d\rho}{\rho} = -\frac{1}{H} dz$$

The indefinite integral of this equation is

$$\int \frac{d\rho}{\rho} = \ln(\rho) = -\frac{1}{H} \int dz = -\frac{z}{H} + \text{a constant}$$

Determining the constant requires further information. This is most commonly in the form of a “boundary condition”. For instance if $\rho = \rho_0$ when $z = 0$, the constant must obviously be $\ln(\rho_0)$. We then have

$$\ln(\rho) = -\frac{z}{H} + \ln(\rho_0)$$

Which can be re-arranged to give

$$\ln(\rho) - \ln(\rho_0) = \ln\left(\frac{\rho}{\rho_0}\right) = -\frac{z}{H}$$

The defining relation for the natural log then implies that

$$\frac{\rho}{\rho_0} = e^{-\frac{z}{H}}$$

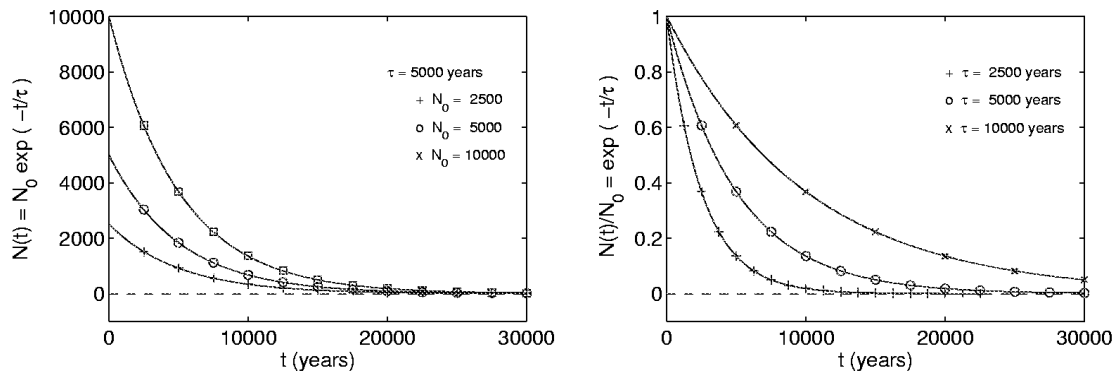
or

$$\rho = \rho_0 e^{-\frac{z}{H}}$$

We will see that this equation gives the exponential decay of the density (mass per unit volume) of a planet's atmosphere as a function of altitude. The constant H is called the "scale height" and is the distance (about 10 km for Earth) over which the density decreases by a factor of $e^{-1} = 0.37$. The relationship of H to the physical properties of the atmosphere and the planet's gravity will be described later.

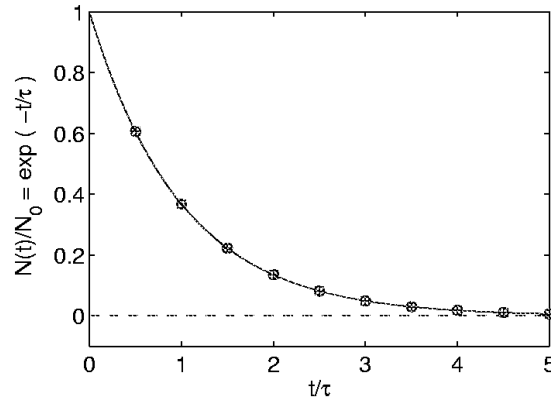
Exponentials with real arguments

Here are two plots of the decaying exponential function $N_0 e^{-t/\tau}$:



If t is time, then τ is called the "time constant" or "decay time". It is the time for the function to decay to $e^{-1} = 37\%$ of its original value. In two time constants, the function has decayed to 13.5%; in three time constants it has decayed to 5% and in 5 time constants it is less than 1%. Discrete points are plotted on each curve at time intervals of $\tau/2$. The plot on the left shows the function in its full "dimensional" form in which the actual value of the function is plotted versus time. N_0 is the starting value of the function. It might, for instance, be the initial number of radioactive Carbon 14 atoms in a sample. The curves on the right are normalized by N_0 . The three curves on the left then coalesce into the single curve for $\tau = 5,000$ years on the right. The shape of this curve does not depend on the initial value. The two other curves on the right are for other values of τ . An

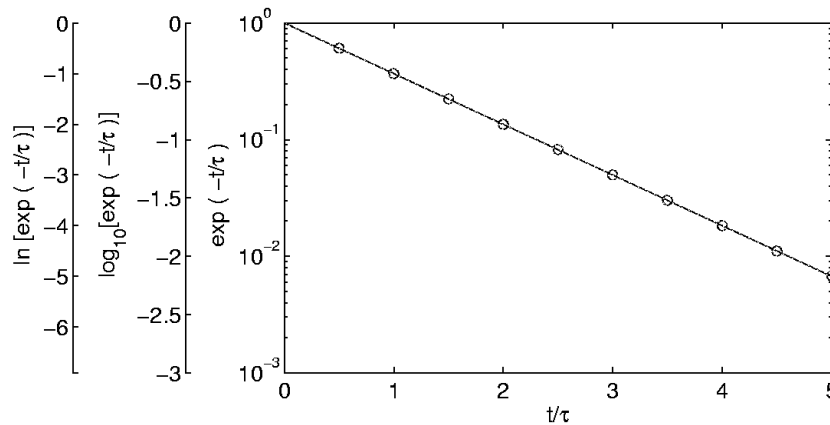
even more useful way to plot the same data is its completely “non-dimensional” form in which the time axis is normalized by the decay time. Then the three curves on the right above coalesce to the single curve plotted in this figure:



Because exponentials become small so rapidly, it is difficult to see what is happening in the above plot after several time constants. In particular, deviations from exponential behavior would be hard to detect. “Logarithmic” plots are a useful alternative. If we take the natural logarithm of $e^{-t/\tau}$ we obtain:

$$\ln(e^{-t/\tau}) = -t/\tau$$

This next plot of this relation gives a straight line with a slope of -1 :



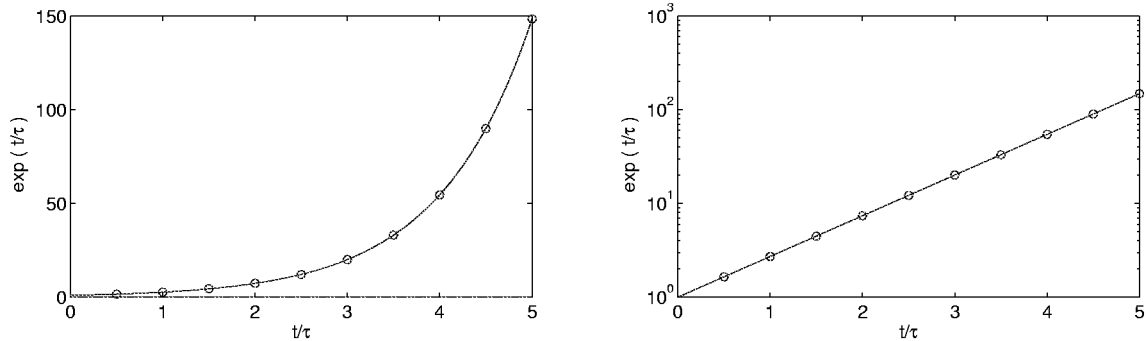
The left-hand scale uses the natural log values. We rarely actually plot the natural log, however. Instead, we usually take the base 10 logarithm giving

$$\log_{10}(e^{-t/\tau}) = -(t/\tau) \log_{10} e = -0.4343 (t/\tau)$$

A plot of the base 10 log of an exponential is still a straight line, but the \log_{10} scale shown in the middle has been “stretched” and thus the slope is -0.4343 instead of -1 . Finally, instead of using the \log_{10} values as tick mark labels on the vertical axis, it is common to use the values of the function itself. Thus -2 is labeled 10^{-2} etc. This “logarithmic scale” is shown to the right. This type of graph is called “semi-logarithmic” or simply semi-log

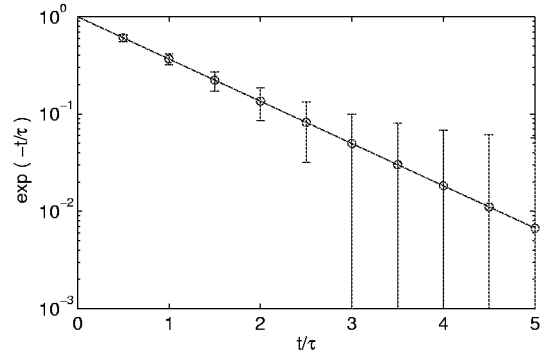
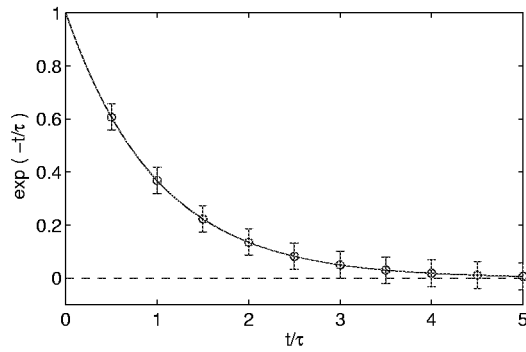
because only one of the axes is logarithmic. These plots can be automatically generated by most graphing software or can be plotted by hand on semi-log graph paper.

A semi-log plot is equally useful when the behavior is a growing exponential:



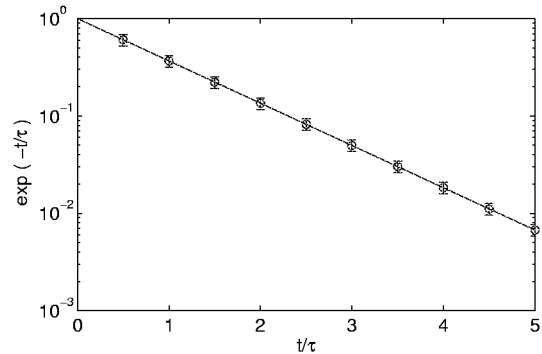
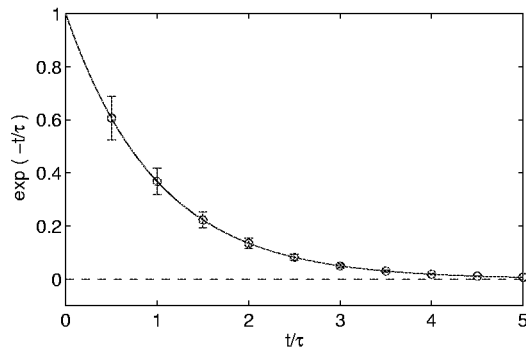
The radioactive decay of Carbon isotope 14 has a time constant of 5730 years. This decay time is determined by counting the number of energetic particles ejected in a given time by the break-up of C^{14} . Carbon isotope 12, on the other hand, is stable and does not decay. The ratio of the number of atoms of C^{14} to C^{12} in the atmosphere is kept nearly constant by the influx of cosmic rays. Living organisms are constantly exchanging carbon with the atmosphere and thus have a C^{14} to C^{12} ratio equal to the atmosphere. When the organism dies, however, it stops exchanging carbon with the atmosphere and the ratio decays with a time constant equal to the decay of C^{14} . By measuring the C^{14} to C^{12} ratio in dead trees killed when their roots were inundated by salt water during subsidence at the time of a large, pre-historic earthquake, one can determine when an earthquake occurred. This is how we know that the last great earthquake along the Pacific Northwest coast was about 300 years ago.

However, measurements without error are not possible. The next plot below on the left is the standard decay curve with “absolute” error bars of ± 0.05 added to “data” at time intervals of 0.5τ . The semi-log equivalent is shown on the right. First note that we would not be able to distinguish the value of the function from 0 for $t/\tau \geq 3$. Thus, if this was the level of error for Carbon 14 dating, we could only date material with an age of less than 3 time constants or about 17,000 years. Decreasing the measurement error helps, but not much. Reducing absolute error a factor of 10 to ± 0.005 extends the maximum age that can be determined to only about 5 time constants; another factor of 100 reduction to ± 0.00005 gets you to about 10 time constants. Modern instrumentation can achieve very small error levels, but Carbon 14 dating is unlikely to ever exceed 100,000 years.



Note that, while the error bars have equal length on the plot of the decaying exponential to the left, the same error bars are not of equal length on the semi-log plot to the right.

In the two plots below, the error bars for all have the same “relative” or “percentage” errors. They have been scaled so that the bars have at $t=\tau$ are the same length as the absolute errors above. In the “absolute” error plots above have the same length a . Because the relative error is a fixed percentage of the value of the function (13.5%), the lengths of the bars decrease as the function decreases and the error bars on the semi-log plot are all equal.



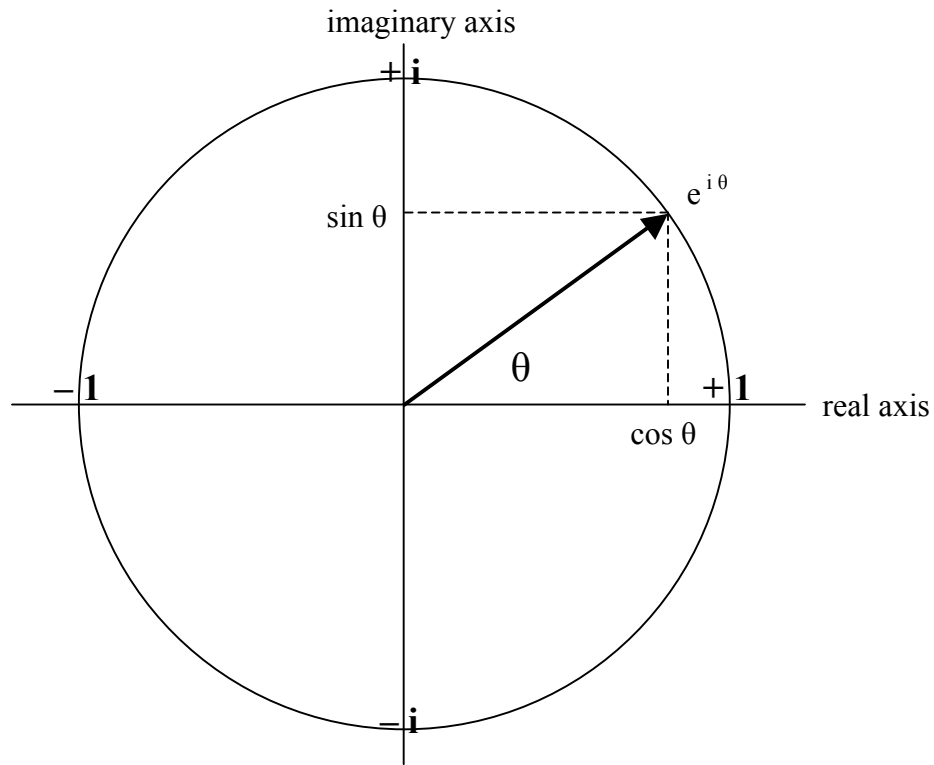
It is important to understand the difference in behavior of “absolute” and “relative” error bars. In particular, if you encounter error bars that are all the same size on a semi-log plot, it should awaken a healthy skepticism. They imply that measurements are more accurate when the “signal” is smaller. This is a highly unlikely state of affairs.

Exponentials with imaginary arguments.

The exponential function $e^{i\theta}$ is complex. Its magnitude squared is

$$|e^{i\theta}|^2 = e^{i\theta} (e^{i\theta})^* = e^{i\theta} e^{-i\theta} = e^{i\theta - i\theta} = e^0 = 1$$

Plotted on the complex plane it is a vector of unit length pointing away from the origin and making an angle θ with respect to the real axis. The projection of this vector on the real axis is $\cos \theta$ and its projection on the imaginary axis is $\sin \theta$.



Thus the imaginary exponential can be written in its complex form as:

$$e^{i\theta} = \cos \theta + i \sin \theta$$

and

$$e^{-i\theta} = \cos \theta - i \sin \theta$$

From these two equations, the following expressions are easily derived:

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

The imaginary part of the argument $i\theta$ (i.e. θ) is called the “phase”. Commonly the phase is replaced by kx for oscillations in space or ωt for oscillations in time. The “wavenumber” k has units of radians per meter. It is related to the “wavelength” λ of a complete oscillation of either the real (cosine) or imaginary (sine) part of the oscillation by

$$k = \frac{2\pi}{\lambda}$$

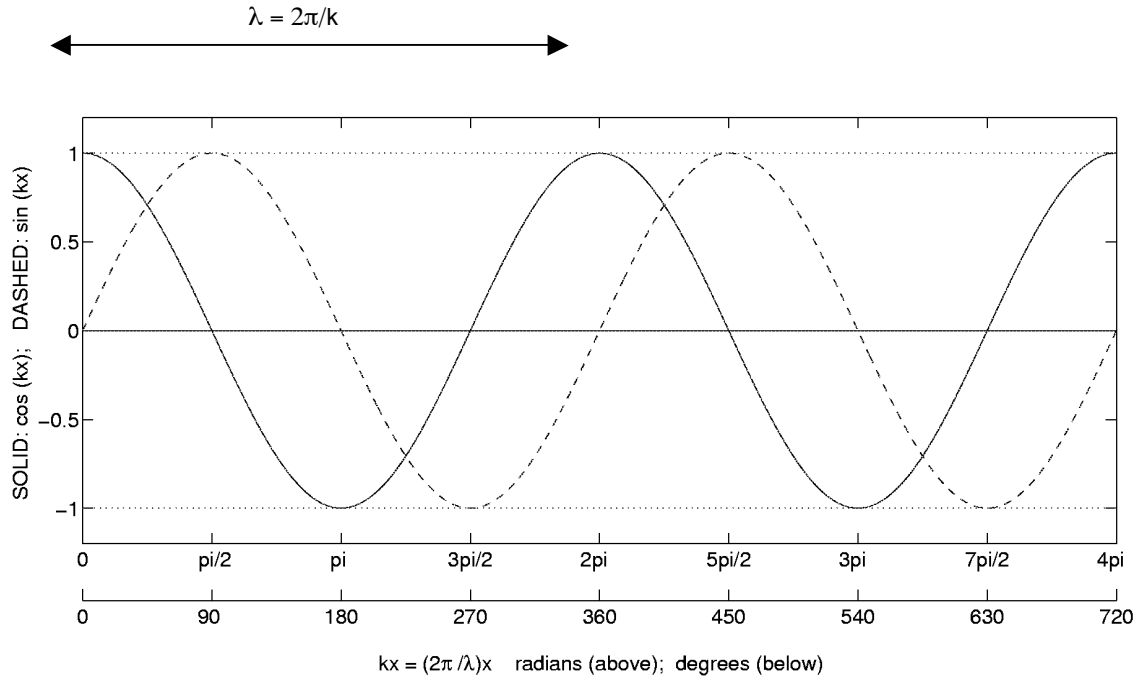
The “angular frequency” ω is related to the frequency for a complete oscillation by

$$\omega = 2\pi f$$

and to the period T of a complete oscillation by

$$\omega = \frac{2\pi}{T}$$

Here is a plot of two complete oscillations of the real (cosine) and imaginary (sine) parts as a function of kx .



Note that the sine function is identical to the cosine function except for being shifted to the right by $\pi/2$ radians (90 degrees).

Differentiation of a sinusoidal oscillation does not change its shape. Instead, it shifts its phase to the left by $\pi/2$. This is most easily seen for the sine. The derivative of $e^{i\theta}$ is

$$\frac{d}{d\theta} e^{i\theta} = i e^{i\theta} = i (\cos \theta + i \sin \theta) = -\sin \theta + i \cos \theta$$

As noted earlier, when complex quantities are equal, their real parts and their imaginary parts must be separately equal. Thus

$$\frac{d}{d\theta} \sin \theta = \text{Im} \left\{ \frac{d}{d\theta} e^{i\theta} \right\} = \cos \theta$$

So the derivative of the dashed curve above is the solid curve. Likewise

$$\frac{d}{d\theta} \cos \theta = \operatorname{Re} \left\{ \frac{d}{d\theta} e^{i\theta} \right\} = -\sin \theta$$

which is also simply a cosine with a $\pi/2$ negative phase shift as you can see by mentally flipping the dashed curve above upside down

Exponentials and complex numbers

It should be obvious from the definition of the complex plane and the discussion of an exponential with an imaginary argument, that a general complex number in its rectangular form can also be written in its “polar” form

$$a + ib = re^{i\theta}$$

where

$$r = |a + ib| = \sqrt{a^2 + b^2}$$

is the distance from the origin to the point representing the complex number on the complex plane and

$$\theta = \tan^{-1} \left(\frac{b}{a} \right)$$

is the phase of the complex number. When making this transformation, you need to be careful that the phase is in the correct quadrant of the complex plane.

Traveling sinusoidal waves

We have considered only functions of a single variable. Waves in nature, however, depend on both distance and time. We shall show later that the height of a cosine disturbance on the surface of an ocean (or your bathtub) can be written

$$h(x, t) = A \cos(kx - \omega t)$$

where A is the deviation of the crest of the wave from the wave’s average height. At a fixed time (such as $t = 0$), this function is the solid curve plotted above. What you see is a surface that undulates spatially with a wavelength $\lambda = 2\pi/k$. If, however, you concentrate on a fixed point in space (such as $x = 0$), what you see is a surface that oscillates up and down with a period $\tau = 2\pi/\omega$. Now suppose you want to always be at a crest of the wave, you would need the phase of the cosine in the last expression above to be $0, 2\pi, 4\pi$, etc. So you would need

$$kx - \omega t = 2n\pi \quad n = 0, 1, 2 \dots$$

which can be rearranged as

$$x = \left(\frac{\omega}{k} \right) t + \frac{2n\pi}{k}$$

Taking the derivative of this with respect to t gives

$$\frac{dx}{dt} = \frac{\omega}{k} = c_p$$

This is the velocity at which you need to move to stay on the crest. This velocity is called the “phase velocity” and it will always be denoted c_p . You can easily repeat this argument using the trough of the wave or, for that matter, any wave phase. Note also that the wave above has a positive phase velocity (moves in the direction of positive x). The wave

$$h(x, t) = A \cos(kx + \omega t)$$

moves in the negative x direction.

Solving another simple ODE and finding the extreme value of a function

If velocity v is the rate of change of distance x with time t and acceleration a is the rate of change of velocity with time, we can write mathematically

$$v = \frac{dx}{dt}$$

$$a = \frac{dv}{dt} = \frac{d}{dt} \left(\frac{dx}{dt} \right) = \frac{d^2x}{dt^2}$$

Suppose we want to understand the dynamics of water that is erupted from a geyser such as “Old Faithful” in Yellowstone National Park. Once the water is squirted out of the “throat” of the geyser, it is acted on only by Earth’s gravity. Thus each “particle” of water feels the “acceleration of gravity”. We express this mathematically as

$$\frac{d^2z}{dt^2} = -g$$

This is also very simple ODE of “2nd degree” that comes up over and over again. The minus sign on the right is because we have defined vertical position z to be positive upwards and the acceleration of gravity is downwards. We will let the ground surface (i.e. the exit of the throat) be $z = 0$ and measure time from the instant that the water leaves the throat. The indefinite integral of this equation is

$$\frac{dz}{dt} = -gt + A$$

where A is the integration constant. A useful boundary condition to constrain A is the velocity v_0 of the water at the instant it leaves the geyser throat. Thus at $t = 0$

$$\frac{dz}{dt} = A = v_0$$

and therefore the water velocity at times greater than 0 is

$$v(t) = \frac{dz}{dt} = -gt + v_0$$

However, measuring v_0 is not an easy task and so this result is not very useful. We can, however, solve for v_0 in terms of a quantity much easier to observe: the maximum height of the eruption H . This last equation above is an ODE of 1st degree and its indefinite integral is

$$z = -\frac{1}{2}gt^2 + v_0t + C$$

where C is another integration constant. Since $z = 0$ at $t=0$, we must have $C = 0$ and so

$$z(t) = -\frac{1}{2}gt^2 + v_0t$$

At this point, we could find v_0 by brute force: We could guess its value and compute $z(t)$ for many values of t to find the $z(t)$ that is maximum. If this is not equal to H , we could keep altering our guess for v_0 until we do get H . Calculus gives us a more elegant and much easier way: Our expression for $v(t)$ states that the water velocity decreases with time as the water rises. This continues until the velocity drops to 0 after which the water starts to descend. Thus the top of the eruptive water column (the maximum value of z) occurs when

$$\frac{dz}{dt} = v = 0$$

The derivative of our expression for $z(t)$ (which is our expression for $v(t)$) implies that at the maximum water altitude

$$0 = -gt + v_0$$

which can be solved to give

$$t = \frac{v_0}{g}$$

Substituting this into $z(t)$ we get

$$H = -\frac{g}{2}\left(\frac{v_0}{g}\right)^2 + v_0\left(\frac{v_0}{g}\right) = \frac{v_0^2}{2g}$$

Which can be re-arranged to give

$$v_0 = \sqrt{2gH}$$

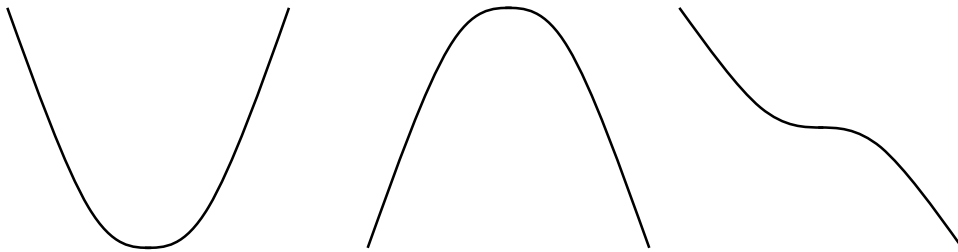
Thus if $H = 50$ m and $g = 10$ m/s², $v_0 = 32$ m/s = 110 km/h. We have learned something useful by finding the extreme value of a function.

In general, the maximum or minimum of a function $f(x)$ occurs for the x at which

$$\frac{df}{dx} = 0$$

This equation may have more than one value of x at which it is true and so there may be more than one “extreme point”. If you are looking for the global (i.e. biggest) maximum (or smallest minimum) you will have to examine all of the extreme points.

You can determine whether you have found a maximum or minimum by looking at the 2nd derivative. All three curves in the next figure have a point at which the slope is 0.



The curve that is concave upward has a minimum. Its slope is negative to the left of the minimum, passes through 0 at the minimum and becomes increasingly positive to the right. Thus the rate of change of the slope (i.e. the 2nd derivative) is positive at the minimum. The curve that is concave downward has a maximum. Its slope is positive to the left of the maximum, passes through 0 at the maximum and becomes increasingly negative to the right. Thus its rate of change of the slope (i.e. the 2nd derivative) is negative at the minimum. For the geyser eruption, the 2nd derivative of $z(t)$ is the gravitational acceleration and is indeed negative as we expect for a maximum. The third curve has a point called an “inflection” where the slope is zero but which is neither a maximum nor a minimum. Instead the concavity of the function changes from upward to downward and the 2nd derivative is zero at this inflection point. Thus you can easily distinguish an inflection point from a maximum or minimum..

Taylor series

This is one of the most important relationships in mathematical physics. For a general function $f(x)$ one can approximate its value in the neighbourhood of $x = a$, by a polynomial series:

$$f(x) \approx f(a) + \frac{f'(a) \cdot (x-a)}{1} + \frac{f''(a) \cdot (x-a)^2}{2!} + \dots + \frac{f^n \cdot (x-a)^n}{n!}.$$

Where f' stands for the df/dx , and f^n stands for $d^n f/dx^n$. Oftentimes, we are in the situation where x is sufficiently close to a that we can make say $(x-a)^2 \ll (x-a)$, and so neglect all second order terms and higher. The above equation can be written as

$$f(x) \approx f(a) + f'(a) \cdot (x-a) + O(x-a)^2$$

where $O(x-a)^2$ is a common way of writing “terms of order $(x-a)^2$ and higher”, which we then neglect. After which we can write

$$f'(a) \approx \frac{f(x)-f(a)}{(x-a)},$$

which is familiar to you as the approximation of the slope of the curve at $x = a$.

Note that another way of writing the Taylor series expansion is

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \frac{f''(x)\Delta x^2}{2!} + \dots + \frac{f^n(x)\Delta x^n}{n!}$$

Applications

1. Small number expansions.

Taylor series can be used, for instance, to calculate approximate values for small numbers without needing a calculator.

Example: What is the value of $(4.0004)^{\frac{1}{2}}$?

Answer: Let $f(x) = x^{\frac{1}{2}}$, and let $a = 4.0$. In which case $f' = \frac{1}{2}x^{-\frac{1}{2}}$, and from the above approximation we can write that:

$$f(4.0004) \approx 4^{\frac{1}{2}} + \frac{1}{2}(4)^{-\frac{1}{2}} \cdot (4.0004 - 4) + O(4.0004 - 4)^2$$

Neglecting the second order terms this gives

$$f(4.0004) \approx 2 + \frac{1}{4}4 \times 10^{-4} \approx 2.0001$$

This is accurate to about 8 significant figures!

2. Taylor series and conservation laws

Another place we will see Taylor series over and over is in developing differential equations. Here is an example for conservation of mass.

Let the density (i.e., mass per unit volume) be a function of space and time: $\rho = \rho(x, y, z, t)$. Now, at a completely arbitrary point in space, consider a small cube of size $\Delta x \times \Delta y \times \Delta z$. The mass of the cube is therefore $\Delta m = \rho \Delta x \Delta y \Delta z$.

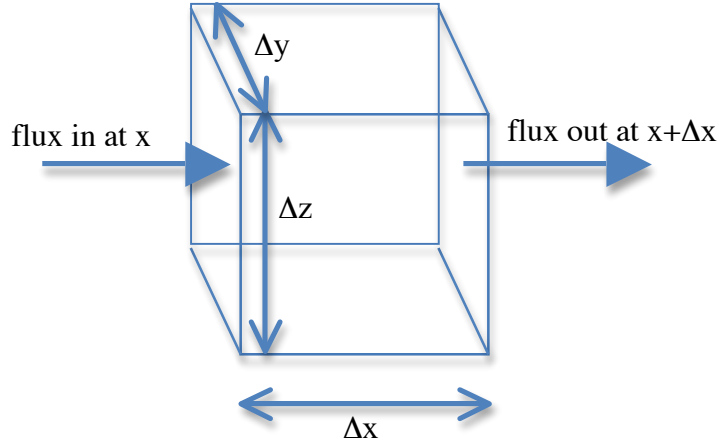


Figure for using Taylor series for deriving conservation laws

How does the density of the cube change with time?

First, consider mass flows in the x direction only. The mass flux in the x direction is equal to the ρu , where $u = dx/dt$, the speed of mass flow in the x direction.

Conservation of mass requires that the rate of change of mass with time is equal to the total flow of mass into the cube minus the total flow of mass out:

$$\frac{d}{dt} \Delta m = \text{total mass flow in} - \text{total mass flow out.}$$

The total mass flow is the flux times the area normal to the flow ($= \rho u \Delta y \Delta z$). The above equation can be written

$$\frac{d}{dt} (\rho \Delta x \Delta y \Delta z) = \Delta y \Delta z (\rho u|_x - \rho u|_{x+\Delta x})$$

where the $\Delta y \Delta z$ is common to both terms on the right hand side, the expression for Δm has been substituted in; and the notation $\rho u|_x$ means the value of ρu calculated at x . Here is where the Taylor series comes in. We can write:

$$\rho u|_{x+\Delta x} \approx \rho u|_x + \frac{d}{dx} (\rho u|_x) \Delta x + O(\Delta x^2)$$

Neglecting the second order terms, this can be substituted into the mass conservation equation to give

$$\frac{d\rho}{dt} = -\frac{d(\rho u)}{dx}$$

or

$$\frac{d\rho}{dt} + \frac{d(\rho u)}{dx} = 0$$

The result can be generalized to flow in all three directions and written as

$$\frac{d\rho}{dt} + \nabla \cdot \rho \mathbf{u} = 0$$

where \mathbf{u} is the three dimensional velocity vector, and $\nabla \cdot ()$ is the divergence operator. The above equation is the *vector equation for conservation of mass*, and has lots of applications.

Functions with more than one argument: Partial derivatives and the gradient

Differentiation can readily be generalized to functions with more than one variable. A “partial derivative” of a function is its rate of change when we make a change in only one of the function arguments. The symbol ∂ replaces d in the derivative operator to remind us that we are only varying one of the arguments. Consider the function

$$h(x, y) = e^{-(x^2+y^2)}$$

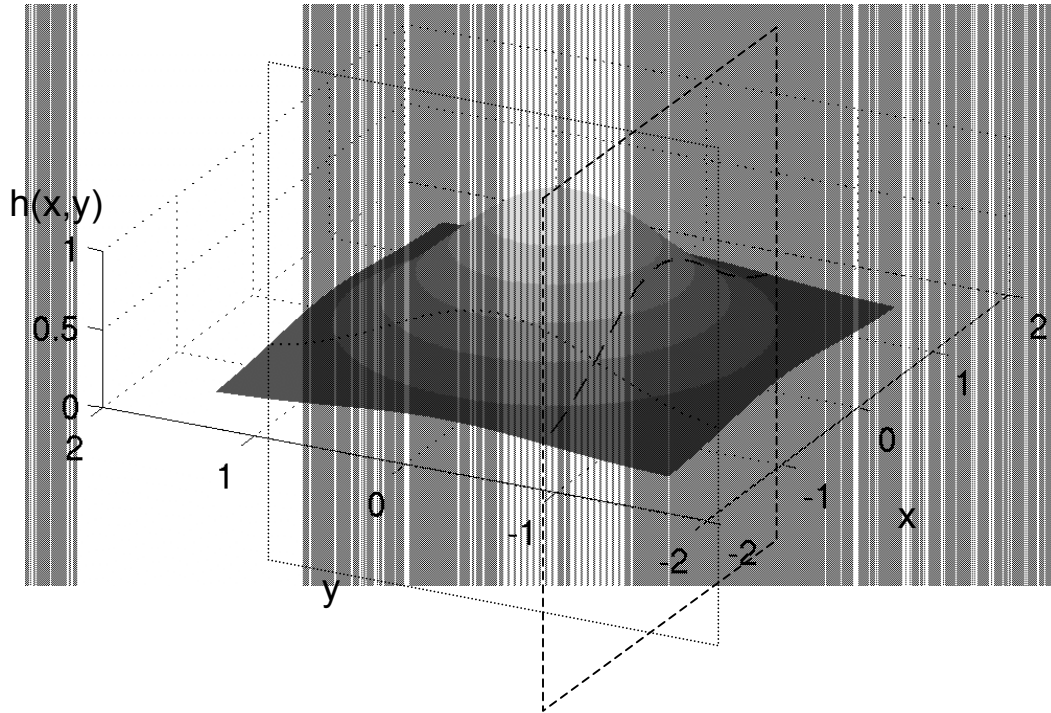
There are two partial derivatives

$$\frac{\partial}{\partial x} h(x, y) = e^{-(x^2+y^2)} \frac{\partial}{\partial x} [-(x^2 + y^2)] = -2x e^{-(x^2+y^2)}$$

and

$$\frac{\partial}{\partial y} h(x, y) = e^{-(x^2+y^2)} \frac{\partial}{\partial y} [-(x^2 + y^2)] = -2y e^{-(x^2+y^2)}$$

A plot of $h(x,y)$ as a three-dimensional (3D) surface is a hill symmetric about the origin $(x, y) = (0, 0)$:



Geometrically, a fixed value of y in this figure is a vertical plane parallel to the x and h axes. The plane defined by $y = -0.66$ is shown outlined with dashes. This plane intersects the surface of the hill along the dashed curve. Then

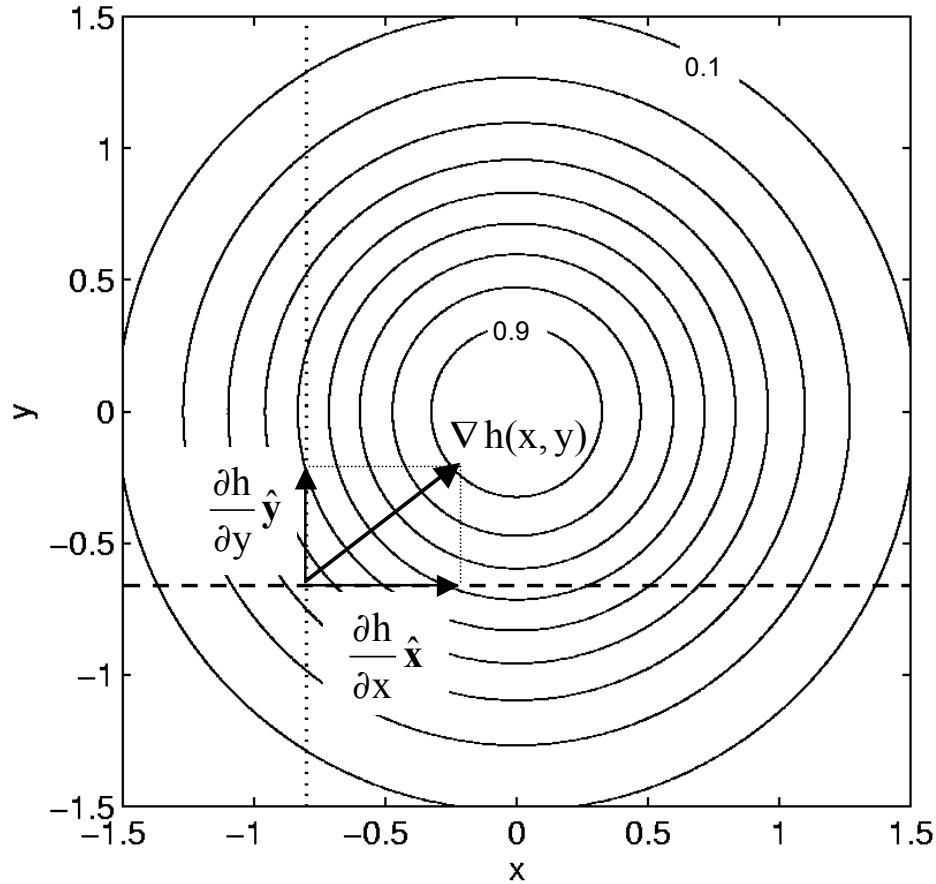
$$\left(\frac{\partial h}{\partial x} \right)_{y=-0.66} = -2x e^{-(0.66)^2} e^{-x^2} = 1.29x e^{-x^2}$$

is the slope of the dashed curve. Likewise, a fixed value of x is a vertical plane parallel to the y and h axes. The plane defined by $x = -0.80$ is shown outlined with dots. It intersects the surface of the hill along the dotted curve and

$$\left(\frac{\partial h}{\partial x} \right)_{x=-0.80} = -2y e^{-(0.8)^2} e^{-y^2} = 1.05y e^{-y^2}$$

is the slope of the dotted curve.

In the above plot, intervals of altitude are shown as regions of different gray. The boundaries between the regions have constant altitude and are called contour lines. Here is a contour plot (or map) of the same function.



Think of this as the hill seen from directly above. The intersections of the two vertical planes with the hill are the straight dashed and dotted lines. At the intersection of these two lines, $(x, y) = (-0.7, -0.66)$, we have drawn the vector

$$\nabla h = \frac{\partial h}{\partial x} \hat{\mathbf{x}} + \frac{\partial h}{\partial y} \hat{\mathbf{y}}$$

This vector is called the “gradient” of h . It is the multi-dimensional equivalent of the derivative. Its components are the slopes of the dashed and dotted curves at their point of intersection. The magnitude of the gradient

$$|\nabla h| = \sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}$$

is the maximum slope of the plane tangent to the $h(x, y)$ surface at the point of intersection. Thus the gradient vector is perpendicular to the contour lines. It points in the uphill direction.

Proof of this result:

The direction of the steepest slope. Starting at some point in horizontal space (x,y) , the elevation at that point is then $h(x,y)$. Now imagine moving away from this starting point by a small distance dx in the x direction and a small distance dy in the y direction. The vector represented this motion is denoted by $\mathbf{dr} = \hat{\mathbf{x}}dx + \hat{\mathbf{y}}dy$. The change in height is given by

$$dh = \left(\frac{dh}{dx}\right)dx + \left(\frac{dh}{dy}\right)dy$$

The steepest slope will be when the direction of \mathbf{dr} is such that dh is a maximum value. The right hand side can be expressed as

$$\left(\frac{dh}{dx}\right)dx + \left(\frac{dh}{dy}\right)dy = \nabla\mathbf{h} \cdot \mathbf{dr}$$

The largest value of the dot product, and hence the largest change is when the two vectors are pointed in the same direction. In other words, the largest change in dh (steepest slope) is when \mathbf{dr} and $\nabla\mathbf{h}$ are parallel. $\nabla\mathbf{h}$ points in the direction of the maximum slope. QED.

The magnitude of the steepest slope. This is the elevation change divided by the horizontal distance travelled:

$$\text{slope} = \frac{dh}{|\mathbf{dr}|} = \nabla\mathbf{h} \cdot \frac{\mathbf{dr}}{|\mathbf{dr}|}$$

The second term in the dot product is, by definition, a unit vector and so has magnitude 1. We have just shown that \mathbf{dr} and $\nabla\mathbf{h}$ are parallel. So the magnitude of the slope is just $|\nabla\mathbf{h}|$. QED.

Maxima and minima in more than one dimension: Least Squares line fitting (statistics, really)

Finding maxima and minima in more than one dimension is an obvious generalization of the case in one dimension: To find the top of the hill in the last example, you go in the direction of the gradient vector until you find a point at which there is no further altitude gain in any direction of motion (the magnitude of the gradient vector is zero). At this point, the tangent plane is horizontal and the partial derivatives of h with respect to x and y must both be zero.

A very important practical example is fitting a function to data by minimizing the misfit between the function and the measurements. We will illustrate this for the case where the function is a straight line and misfit is judged by the square of the “distance” between the line and each data point. However, the principle is valid for more complicated functions and other ways of judging misfit.

Suppose we have measured N data points (x_i, y_i) and that we have a theory that predicts that the relation between x and y is the straight line:

$$y = mx + b$$

We can estimate the slope m and the y -intercept b by finding their values that “fit” the data “best”. To do this, we need to define what we mean by “fit” (more precisely misfit) and what we mean by “best”. The misfit (often called the “residual”) of a data point can be defined as the difference between what we predict from the “theoretical” relation and the value we measured. Thus for the i^{th} point, the misfit is

$$r_i = y_{\text{predicted}}(x_i) - y_i = mx_i + b - y_i$$

A useful measure of the total misfit is the sum of the squared misfits for all the data points

$$Q = \sum_{i=1}^N (r_i)^2 = \sum_{i=1}^N (mx_i + b - y_i)^2$$

We can then define the “best” fit as the values of m and b that make Q smallest. From our earlier discussion, it should be clear this happens when the two equations

$$\frac{\partial Q}{\partial m} = 0$$

and

$$\frac{\partial Q}{\partial b} = 0$$

are simultaneously satisfied. Using the rules for differentiation listed earlier and noting that partial differentiation with respect to m is computed with x_i , y_i and b constant, the first of these equations is

$$\begin{aligned} \frac{\partial}{\partial m} \sum_{i=1}^N (r_i)^2 &= \sum_{i=1}^N \frac{\partial}{\partial m} (r_i)^2 = \sum_{i=1}^N \frac{\partial}{\partial m} (mx_i + b - y_i)^2 \\ &= \sum_{i=1}^N 2(mx_i + b - y_i) \frac{\partial}{\partial m} (mx_i + b - y_i) \\ &= 2 \sum_{i=1}^N (mx_i + b - y_i) x_i = 2 \left[m \sum_{i=1}^N (x_i)^2 + b \sum_{i=1}^N x_i - \sum_{i=1}^N (x_i y_i) \right] = 0 \end{aligned}$$

and the second equation is

$$\begin{aligned}
 \frac{\partial}{\partial b} \sum_{i=1}^N (r_i)^2 &= \sum_{i=1}^N \frac{\partial}{\partial b} (r_i)^2 = \sum_{i=1}^N \frac{\partial}{\partial b} (mx_i + b - y_i)^2 \\
 &= \sum_{i=1}^N 2(mx_i + b - y_i) \frac{\partial}{\partial b} (mx_i + b - y_i) \\
 &= 2 \sum_{i=1}^N (mx_i + b - y_i)(1) = 2 \left[m \sum_{i=1}^N x_i + b \sum_{i=1}^N (1) - \sum_{i=1}^N y_i \right] = 0
 \end{aligned}$$

Defining

$$A = \sum_{i=1}^N (x_i)^2 \quad B = \sum_{i=1}^N x_i \quad C = \sum_{i=1}^N (x_i y_i) \quad D = \sum_{i=1}^N (1) = N \quad E = \sum_{i=1}^N y_i$$

these two equations can be written

$$Am + Bb = C$$

and

$$Bm + Db = E$$

Both must be satisfied, so solving them simultaneously we get

$$m = \frac{DC - BE}{DA - B^2}$$

$$b = \frac{AE - BC}{DA - B^2}$$

1.2 Physical tools

Conservation laws

These have enormous power to explain what we see around us, and a big aspect of the course will be showing why.

Conservation of Energy. Energy cannot be created or destroyed, it can only be converted from one form into another (potential, kinetic, electrical, etc.)

Conservation of Mass. Absent thermonuclear reactions, the same mass of stuff goes in as comes out. Apart from the Sun and dating methods using radioactive decay (why?), this is good enough for us on Earth.

Conservation of linear momentum. Really Newton's laws, see below.

Conservation of angular momentum. A circular version of Newton's laws (see below)

Conservation of electric charge. We won't care about this that much, the total number of electrons and protons stays the same.

(n.b. could put more in here about closed systems)

Newton's laws of motion

1. *An object will continue in a state of rest or of constant motion unless acted upon by an unbalanced force.* Inertia (i.e. mass) keeps you doing what you're doing unless someone pushes you.

2. *The rate of change of momentum of an object is proportional to the unbalanced forces.* Basically, $F=ma$.

3. *Every action has an equal and opposite reaction.* Push me and I'll push back on you.

Formulae governing uniform acceleration

Let an object start at $x = 0$ at time $t = 0$ with initial velocity, u , and let it be subject to a uniform acceleration, a . Its subsequent position, x , and velocity, v , at some later time, t , are described by the following formulae:

$$v^2 = u^2 + 2ax$$

$$x = x_0 + ut + \frac{1}{2} at^2$$

$$v = u + at$$

$$x = \frac{1}{2} (u+v)t$$

Appendix: maybe useful, but not essential

This section presents more advanced materials that build on what is presented above. You may not have seen it before. To the extent that we use it, it will be covered in detail in context. It is provided for the interest of students with more advanced background than is required.

Weighted least squares

The measure of misfit used above is, by no means, the only useful measure. For instance, suppose you want to weight the misfit of each data point in the overall misfit by a different amount w_i . The overall misfit can be re-defined as

$$Q_{weighted} = \sum_{i=1}^N (w_i r_i)^2 = \sum_{i=1}^N [w_i (x_i + b - y_i)]^2$$

Our best fit is fit can then be defined by the equations

$$\frac{\partial Q_{weighted}}{\partial m} = 0$$

and

$$\frac{\partial Q_{weighted}}{\partial b} = 0$$

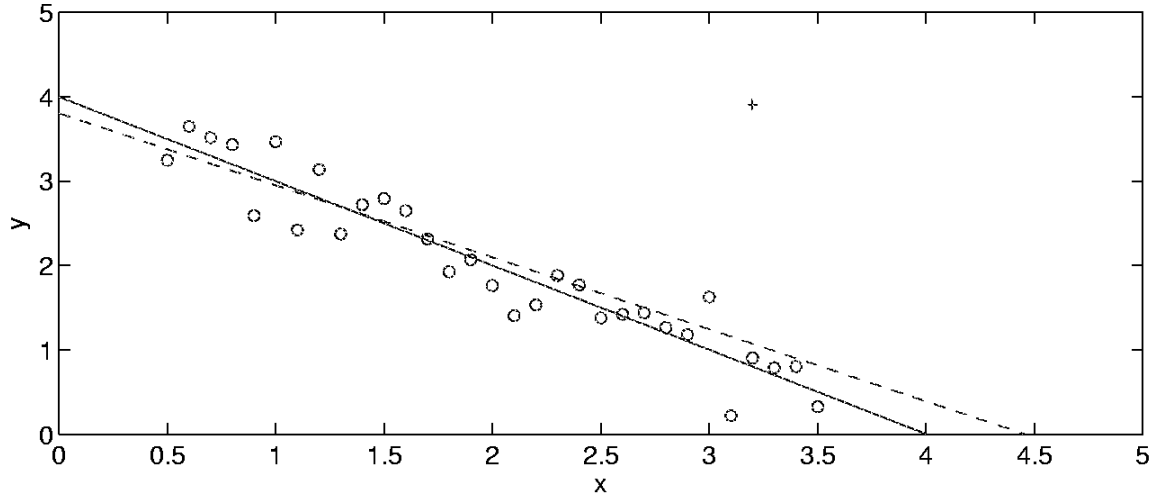
Following through the rest of the algebra leading to the expressions for m and b gives results that are identical except that the constants A to E need to be modified as follows:

$$A = \sum_{i=1}^N (w_i x_i)^2 \quad B = \sum_{i=1}^N (w_i^2 x_i) \quad C = \sum_{i=1}^N (w_i^2 x_i y_i) \quad D = \sum_{i=1}^N (w_i^2) \quad E = \sum_{i=1}^N (w_i^2 y_i)$$

Commonly used weights are

$$w_i = \frac{1}{\sigma_i}$$

where the σ_i are the standard error (square root of the variance) of each data point.



Vector calculus

Because h depends only on x and y , the hill example above involves a two-dimensional gradient. However, the gradient can be generalized to functions that depend on all three spatial scales in an obvious way. The 3D gradient “operator” is

$$\nabla = \hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z}$$

This operator is not a vector, but can be manipulated like a vector. Its dot product with a vector \mathbf{v} is called the “divergence” of \mathbf{v} and can be written in terms of its Cartesian components

$$\nabla \cdot \mathbf{v} = \left(\hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z} \right) \cdot (\hat{\mathbf{x}} v_x + \hat{\mathbf{y}} v_y + \hat{\mathbf{z}} v_z) = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}$$

Note that like the dot product between vectors, the divergence is not a vector. The divergence is related to the volume change of a fluid moving with velocity \mathbf{v} .

The cross product of the gradient operator with the vector \mathbf{v} is called its “curl” and for fluid velocity is related to the angular momentum of the fluid. The cross product can be expressed for a vector with Cartesian coordinates by replacing the components of vector \mathbf{a} in our earlier definition of the cross product with partial derivatives:

$$\nabla \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ b_x & b_y & b_z \end{vmatrix} = \hat{\mathbf{x}} \left(\frac{\partial b_z}{\partial y} - \frac{\partial b_y}{\partial z} \right) - \hat{\mathbf{y}} \left(\frac{\partial b_z}{\partial x} - \frac{\partial b_x}{\partial z} \right) + \hat{\mathbf{z}} \left(\frac{\partial b_y}{\partial x} - \frac{\partial b_x}{\partial y} \right)$$

Useful relations using the gradient operator are:

$$\nabla \times \nabla f = 0$$

$$\mathbf{u} \cdot \nabla f = u_x \frac{\partial f}{\partial x} + u_y \frac{\partial f}{\partial y} + u_z \frac{\partial f}{\partial z}$$

$$\nabla \cdot \nabla f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

which is called the Laplacian.

There are also rules for how the gradient operator works on combinations of functions and vector “fields”. (i.e. vectors that vary in space) that are analogous to the chain rules for ordinary derivatives. Important ones are:

$$\nabla \cdot (f \mathbf{u}) = f \nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla f$$

$$\nabla \times (f \mathbf{u}) = f \nabla \times \mathbf{u} + (\nabla f) \times \mathbf{u}$$

$$\nabla(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot \nabla \mathbf{b} + \mathbf{b} \cdot \nabla \mathbf{a} + \mathbf{a} \times \nabla \times \mathbf{b} + \mathbf{b} \times \nabla \times \mathbf{a}$$

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot \nabla \mathbf{a} - \mathbf{a} \cdot \nabla \mathbf{b} + \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a})$$

Integration can also be easily generalized to multiple dimensions. The integral of a function over a volume V can be approximated by dividing the volume up into small elements, such as cubes, multiplying the average value of the function inside each cube by the volume of the cube and summing up the result for all cubes inside V . The integral over a surface A is similarly approximated by dividing the surface up into small elements, such as squares, multiplying the average value of the function inside each square by the area of the square and summing up the result for all squares inside A . In each case, the true integral is the limit of the sum as you make the elements increasingly small.

A very important relation between volume and surface integrals is called Gauss’ Theorem. If \mathbf{u} is a vector “field” that varies within a closed volume and $\hat{\mathbf{n}}$ is a unit vector that always points out of the volume

$$\oint_{\text{interior of the closed volume}} (\nabla \cdot \mathbf{u}) dV = \oint_{\text{surface of the same volume}} (\hat{\mathbf{n}} \cdot \mathbf{u}) dA$$

The little circle on top of the integral sign is just a reminder that the integration is over the complete surface.