Topics

Daniel Silva, Quintin Mabanta

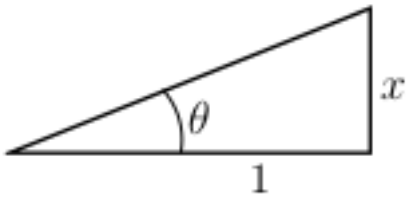
June 2019

Contents

1 The Basics

1.1 Trigonometry

The common trigonometric functions are familiar to you, but do you know some of the tricks to remember (or to derive quickly) the common identities among them? Given the sine of an angle, what is its tangent? Given its tangent, what is its cosine? All of these simple but occasionally useful relations can be derived in about two seconds if you understand the idea behind one picture. Suppose for example that you know the tangent of θ, what is sin θ? Draw a right triangle and designate the tangent of θ as x , so you can draw a triangle with tanθ = x/1.



The Pythagorean theorem says that the third side is $\sqrt{1+x^2}$. You now read the sine from the triangle as x/$\sqrt{1+x^2}$, so

$$\sin\theta = \frac{\tan\theta}{\sqrt{1+\tan^2\theta}} \tag{1.1}$$

Any other such relation is done the same way. You know the cosine, so what's the cotangent? Draw a different triangle where the cosine is x/1.

1.1.1 Hyperbolic Functions

The hyperbolic functions are most intuitively related to the exponential functions.

$$\cosh x = \frac{e^x + e^{-x}}{2} \qquad\qquad \sinh x = \frac{e^x - e^{-x}}{2} \tag{1.2}$$

You can extrapolate the corresponding identities algebraically. E.g.,

$$\cosh^2\theta - \sinh^2\theta = 1 \tag{1.3}$$

The relationships between sine and cosine with their hyperbolic counterparts you can derive from above (or simply from Euler's formula)

$$\sin(ix) = i\sinh(x) \qquad\qquad \cos(ix) = \cosh(x) \tag{1.4}$$

Where do hyperbolic functions occur? If you have a mass in equilibrium, the total force on it is zero. If it's in *stable* equilibrium then if you push it a little to one side and release it, the force will push it back to the center. If it is *unstable* then when you perturb it, the system will tend away from the equilibrium point. In the former case, it will oscillate about the equilibrium position and for small oscillations the function of time will be a circular trigonometric function – the common sines or cosines as a function of time. However, if the point is unstable, the motion will be described by hyperbolic functions. This is obvious from their relationships with the exponential function: one is familiar with the rapid expansion of a real valued exponential and the oscillatory motion of an imaginary valued exponential.

1.2 Integration Techniques

Integrals can be solved via several different methods (not even counting complex methods!). Here we will go over some of the basic ways to approach an ostensibly tough integrand and when some techniques may be more useful than others.

1.2.1 Substitution

The general use of a substitution is taking some complicated integrand and folding some of its difficulty into the infinitesimal differential. The general concept is

$$\int f(g(x))g'(x)dx = \int f(u)du \qquad \text{where} \qquad u = g(x) \tag{1.5}$$

The substitution you choose to make will obviously depend on the integral you're trying to solve; gaining an intuition for what needs to be simplified will come with practice. A simple one might look like

$$\int \left(1 - \frac{1}{x}\right)\cos(x - \ln x)dx \tag{1.6}$$

If we take the function inside the cosine to be some $u = f(x)$, the equation has an obvious form

$$\begin{aligned}\int \frac{du}{dx}\cos(u)dx &= \int \cos(u)du \\ &= \sin(u) + C \\ &= \sin(x - \ln x) + C\end{aligned}$$

For definite integrals, one must be careful to treat the bounds correctly when making these substitutions. When you change the infinitesimal you're integrating along, the bounds are very rarely the same. E.g.

$$\int_{-\pi}^{\frac{\pi}{2}} \cos(x)\cos(\sin(x))dx$$

Let $u = \sin(x)$; the bounds now go from $\sin(-\pi)$ to $\sin(\pi/2)$ and the new integral becomes

$$\begin{aligned} &= \int_0^1 \cos(u)du \\ &= \sin(u)\Big|_0^1 \\ &= \sin(1) \end{aligned}$$

So as long as we're being careful with our bounds, we can make as many substitutions as we like. Here is one last example where we take more than one substitution.

$$\int \frac{1}{x\sqrt{\ln^2 x + 2}}dx$$

Let $u = \ln x$ and we get the following

$$\int \frac{1}{x\sqrt{\ln^2 x + 2}}dx = \int \frac{du}{\sqrt{u^2+2}}$$

Now that this is in a tractable form let $u = \sqrt{2}\tan\theta$. This yields

$$\begin{aligned} &= \int \frac{\sqrt{2}\sec^2\theta}{\sqrt{2}\sec\theta}d\theta \\ &= \int \sec\theta d\theta \\ &= \ln|\sec\theta + \tan\theta| + C \\ &= \ln\left|\frac{\sqrt{u^2+2}}{\sqrt{2}} + \frac{u}{\sqrt{2}}\right| + C \\ &= \ln|\sqrt{\ln^2 x + 2} + \ln x| + C \end{aligned}$$

1.2.2 Integration by Parts

Integrating by parts is one of the brute force methods of integration that one resorts to when there is no obviously elegant way to solve the problem. Viz.

you will be using it a lot (as is the case in a non-elegant universe; sorry Brian Greene!). This technique is especially useful when there is a term that, when hit with a derivative, eventually vanishes. The general form of this technique is found by taking the chain rule of a product of functions, integrating, and applying the fundamental theorem of calculus.

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx \tag{1.7}$$

So you effectively want to denote a part of the integrand to be integrated, and another part to take the derivative of. The most simple example is

$$\int xe^x dx \tag{1.8}$$

If we use the notation from 1.7, we can set $f(x) = x$ and $g'(x) = e^x$. Our equation now becomes

$$\int xe^x dx = xe^x - \int e^x dx = e^x(x-1) + C \tag{1.9}$$

Simple enough but let's look at another example.

$$\int \arctan x dx \tag{1.10}$$

Always be conscious of the trigonometric identities available to you. If we remember that the derivative of $\arctan x$ is $1/(x^2+1)$ this becomes trivial.

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

1.2.3 Partial Fractions

Partial fraction decomposition is a way of taking an integrand and factoring it into simpler expressions, breaking the integral into a sum of integrals. The general idea is: you want to represent your rational expression as some function

$$f(x) = \frac{P(x)}{Q(x)} \tag{1.11}$$

where $P(x)$ and $Q(x)$ are both polynomials (and the former is of smaller degree). Let's take some integral

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

The denominator is easily factorized so let's partial fraction decompose.

$$\begin{aligned}\frac{3x+11}{x^2-x-6} &= \frac{A}{x-3}+\frac{B}{x+2}\\ &= \frac{A(x+2)+B(x-3)}{(x-3)(x+2)}\end{aligned}$$

Now this reduces to simple algebra! We solve the equation

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

By inspection, one can deduce the values for A and B to plug back into our equation

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

Again, this technique can be used for an arbitrarily large degree of polynomial, as long as the numerator has a smaller degree than the denominator.

1.2.4 Completing the Square

This method is typically used when substitution or partial fractions won't work (again, there are several ways to solve an integral so you may find yourself having to try different methods before finding something that sticks). The term "completing the square" essentially means to add zero (in the sense that you're adding some number and taking away the same number) in the hope to factor a quadratic into a square. Let's take a simple example.

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

So we've subtracted 4 from the 5 and lumped it into the left two terms to "complete the square". If we use the same trig identity as before, we can make the substitution $x+2=\tan\theta$ and get

$$\begin{aligned}\int \sqrt{(x+2)^2+1}dx &= \int \sqrt{\tan^2\theta+1}\sec^2\theta d\theta\\ &= \int \sqrt{\sec^2\theta}\sec^2\theta d\theta\\ &= \int \sec^3\theta d\theta\\ &= \frac{1}{2}(\sec\theta\tan\theta+\ln|\sec\theta+\tan\theta|)+C\end{aligned}$$

Substitute back in and you finally get

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

By completing the square we are able to take an integral that has a general quadratic in it and convert it into a form that allows us to use a known integration technique.

1.3 Parametric Differentiation

A useful application of Leibniz's Rule is *differentiating under the integral*. Let's take some example function

$$f(x,n) = \int_0^\infty x^n e^{-x} dx \qquad \text{where } n \in \mathbb{Z}^+ \tag{1.12}$$

You could integrate by parts n times and solve it the slow way, or instead be clever like Feynman. Consider the integral

$$\int_0^\infty e^{-\alpha x} dx \tag{1.13}$$

This integral itself is easy to solve.

$$\int_0^\infty e^{-\alpha x} dx = \frac{-1}{\alpha} e^{-\alpha x}\Big|_0^\infty = \frac{1}{\alpha} \tag{1.14}$$

Now differentiate this with respect to α,

$$\frac{d}{d\alpha}\int_0^\infty e^{-\alpha x} dx = \frac{d}{d\alpha}\frac{1}{\alpha} = \frac{-1}{\alpha^2} = \int_0^\infty x e^{-\alpha x} dx \tag{1.15}$$

If x is independent of α, we can do this process as many times as we'd like. This results in the simple relation

$$\int_0^\infty x^n e^{-\alpha x} dx = \frac{n!}{\alpha^{n+1}} \tag{1.16}$$

Set $\alpha = 1$ and you see that the original integral is just $n!$. We will see later how this relates to the gamma function.

The idea of this method is to change the original problem into another by introducing a parameter. Then differentiate with respect to that parameter in order to recover the problem that you really want to solve. With a little practice, this will become easier than partial integration.

Exercise

Solve the following definite integral by *differentiating under the integral*

$$\int_0^1 \frac{x^2-1}{\log x} dx$$

1.4 Gaussians, ERF, Gamma Function

1.4.1 Gaussian Integrals

Gaussian integrals are an important class of integrals that show up in kinetic theory, statistical physics, quantum mechanics, and almost any scenario involving an ensemble of things to keep track of.

$$\int dx x^n e^{-\alpha x^n} \tag{1.17}$$

The simplest, most common case is the definite integral from -∞ to +∞. Let's look at a special case where n = 0 and $\alpha = 1$. Denote the integral by I, then

$$I = \int_{-\infty}^{\infty} dx e^{-x^2}, \qquad \text{and} \qquad I^2 = \left(\int_{-\infty}^{\infty} dx e^{-x^2}\right)\left(\int_{-\infty}^{\infty} dy e^{-y^2}\right)$$

This can be rearranged to look like

$$I^2 = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy e^{-(x^2+y^2)}$$

Switching to polar coordinates, (How this is done will be elaborated on in 3.6), we arrive at

$$I^2 = \int_0^{2\pi} d\phi \int_0^{\infty} r e^{-r^2} dr$$

The ϕ integral simply gives 2π and the r integral is evaluated to be 1/2. Therefore, our I integral is simply equal to $\sqrt{\pi}$. The same concept outlined before can be applied for the case of different Gaussian integrals.

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}} \quad \text{hence} \quad \int_{-\infty}^{\infty} dx x^2 e^{-\alpha x^2} = \frac{d}{d\alpha}\sqrt{\frac{\pi}{\alpha}} = \frac{1}{2}\left(\frac{\sqrt{\pi}}{\alpha^{3/2}}\right)$$

You can now get the results for higher even powers of x by further differentiating w.r.t. α.

1.4.2 ERF

What about the same integral, but with other limits? The odd-n case is easy to do in just the same way as when the limits are zero and infinity with a simple substitution. The even-n case is different because it can't be done in terms of elementary functions. It is used to define an entirely new function.

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dt e^{-t^2} \tag{1.18}$$

This is called the error function. What can you do with this function? The most likely application is probably to probability. If you flip a coin 1000 times,

you expect it to come up heads about 500 times. But just how close to 500 will it be? If you flip it twice, you wouldn't be surprised to see two heads or two tails, in fact the equally likely possibilities are

TT HT TH HH

This says that in 1 out of $2^2 = 4$ such experiments you expect to see two heads and in 1 out of 4 you expect two tails. For just 2 out of 4 times you do the double flip do you expect exactly one head. *All this is an average. You have to try the experiment many times to see your expectation verified, and then only by averaging many experiments.*

It's easier to visualize the counting if you flip N coins at once and see how they come up. The number of coins that come up heads won't always be $N/2$, but it should be close. If you repeat the process, flipping N coins again and again, you get a distribution of numbers of heads that will vary around $N/2$ in a characteristic pattern. The result is that the fraction of the time it will come up with k heads and $N - k$ tails is, to a good approximation

$$\sqrt{\frac{2}{\pi N}} e^{-2\delta^2/N}, \qquad \text{where} \qquad \delta = k - \frac{N}{2} \tag{1.19}$$

Flip N coins, then do it again and again. In what fraction of the trials will the result be between $N/2$ - Δ and $N/2 + \Delta$ heads? This is the sum of the fractions corresponding to $\delta = 0$, $\delta = \pm 1$, ... , $\delta = \pm\Delta$. Because the approximate function is smooth, I can replace this sum with an integral. This substitution becomes more accurate the larger N is.

$$\int_{-\Delta}^{\Delta} d\delta \sqrt{\frac{2}{\pi N}} e^{-2\delta^2/N}$$

Make the substitution $2\delta^2/N = x^2$ and you have

$$\sqrt{\frac{2}{\pi N}} \int_{-\Delta\sqrt{2/N}}^{\Delta\sqrt{2/N}} \sqrt{\frac{N}{2}} dx e^{-x^2} = \text{erf}(\Delta\sqrt{2/N}) \tag{1.20}$$

The error function of one is 0.84, so if $\Delta = \sqrt{N/2}$ then in 84% of the trials heads will come up between $N/2$ - $\sqrt{N/2}$ and $N/2 + \sqrt{N/2}$ times. For $N = 1000$, this is between 478 and 522 heads.

1.4.3 Gamma Function

A related integral worthy of its own name is the Gamma function.

$$\Gamma(x) = \int_0^\infty dt\ t^{x-1} e^{-t} \tag{1.21}$$

In the special case which x is a positive integer, the gamma function takes on a familiar relationship.

$$\Gamma(n) = (n-1)! \tag{1.22}$$

The factorial is not defined if its argument isn't an integer, but the Gamma function is perfectly well defined for any argument as long as the integral converges. One special case is notable: $x = 1/2$.

$$\Gamma(1/2) = \int_0^\infty dt t^{-1/2} e^{-t} = \int_0^\infty 2udu u^{-1} e^{-u^2} = 2\int_0^\infty du e^{-u^2} = \sqrt{\pi} \quad (1.23)$$

I used $t = u^2$ and then the result from the Gaussian integral. You can use parametric differentiation to derive a simple and useful recursion relation.

$$x\Gamma(x) = \Gamma(x+1) \quad (1.24)$$

From this you can get the value of $\Gamma(1.5), \Gamma(2.5)$, etc.

1.5 Infinite Series & Taylor Series

There are a handful of infinite series that you should memorize and should know just as well as you do the multiplication table. The first of these is the geometric series,

$$1 + x + x^2 + x^3 + x^4 + \cdots = \sum_0^\infty x^n = \frac{1}{1-x} \qquad \text{for} \quad |x| < 1 \quad (1.25)$$

It's very easy derive because in this case you can sum the finite form of the series and then take a limit. Write the series out to the term x^N and multiply it by $(1-x)$. The finite sum up to x^N is useful on its own. For example it's what you use to compute the payments on a loan that's been made at some specified interest rate. You use it to find the pattern of light from a diffraction grating.

$$\sum_0^N x^N = \frac{1-x^N}{1-x} \quad (1.26)$$

Some other common series that you need to know are power series for elementary functions:

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \cdots &&= \sum_0^\infty \frac{x^N}{N!} \\
\sin x &= x - \frac{x^3}{3!} + \cdots &&= \sum_0^\infty (-1)^N \frac{x^{2N+1}}{(2N+1)!} \\
\cos x &= 1 - \frac{x^2}{2!} + \cdots &&= \sum_0^\infty (-1)^N \frac{x^{2N}}{(2N)!} \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots &&= \sum_0^\infty (-1)^{N+1} \frac{x^N}{N}
\end{aligned}$$

The generalization to any function (satisfying certain conditions) is as follows. You match the coefficients in the assumed expansion, and get

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \frac{x^3}{3!}f'''(0) + \cdots = \sum_0^\infty \frac{x^N}{N!}\frac{d^N f}{dx^N}\bigg|_0 \tag{1.27}$$

This is the Maclaurin series, a special case of the Taylor series when expanded around the point zero. This can be expanded about any point, a, to take the form

$$f(a) = f(a)+(x-a)f'(a)+\frac{(x-a)^2}{2!}f''(a)+\frac{(x-a)^3}{3!}f'''(a)+\cdots = \sum_0^\infty \frac{(x-a)^N}{N!}\frac{d^N f}{dx^N}\bigg|_a \tag{1.28}$$

This is the generalized Taylor series and is *indispensable* in several physics problems.

What good are infinite series? This is sometimes the way that a new function is introduced and developed, typically by determining a series solution to a new differential equation. This is a tool for the numerical evaluation of functions. This is an essential tool to understand and invent numerical algorithms for integration, differentiation, interpolation, and many other common numerical methods. To understand the behavior of complex-valued functions of a complex variable you will need to understand these series for the case that the variable is a complex number. All the series that I've written above are power series (Taylor series), but there are many other possibilities. E.g. using the Fourier series you can write a function as an infinite series of sines and cosines instead of an infinite series of powers in the case of Taylor.

1.6 Problems

1. Consider the Gaussian distribution

   $$\rho(x) = Ae^{-\lambda(x-a)^2}$$

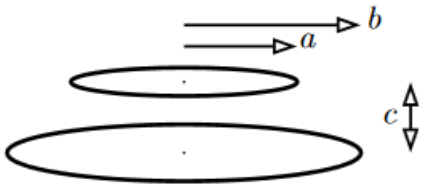
   where A, a, and λ are constants. Assume $\rho(x)$ is a probability distribution. Find the normalization factor, A, in terms of λ and a. Then, find the expectation value of x and x^2.

2. Show that it is possible to stack a pile of identical books so that the top book is as far as you like to the right of the bottom book. You may use as many books as you need, but for any desired finite distance it will always be a finite number.

3. A two-slit diffraction system is illuminated normally with coherent light of wavelength λ. The separation between the slits is a. The intensity of light detected at a distant screen (much further than a), when one of the slits is covered, is I_0. For the situation where both slits are open, calculate and sketch the response of the detector as a function of the angle θ, where θ is measured away from the normal to the system.

4. If there are 100 molecules of a gas bouncing around in a room, about how long will you have to wait to find that all of them are in the left half of the room? Assume that you make a new observation every microsecond and that the observations are independent of each other.

5. Two thin circular rings have radii a and b and carry charges Q_1 and Q_2 distributed uniformly around them. The rings are positioned in two parallel planes a distance c apart and have aligned axes.



Compute the force of one ring on the other. Solve this for the following case: $a = b = 0$ (check that this is consistent with point particles). Also see what happens when $a = b$ and $c \to 0$; you should intuit that the force goes infinity, but perform the integral anyway to see what you arrive at.

2 Linear Algebra

2.1 Determinants

Determinants are mathematical objects that are particularly useful in the analysis and solution of systems of linear equations; but what is the intuition for what a determinant actually is? Start with any region in a plane, and say it has area $\mathcal{A}$. Let's say an operator takes all the vectors which constitute this area and transforms them to create some new area of a size $\mathcal{A}'$. What is the ratio of the new area to the old one? $\mathcal{A}'/\mathcal{A}$. How much does this transformation stretch or squeeze the area? This ratio between these two is indeed the determinant.

The full concept of the determinant allows for negative values. But what would scaling an area by a negative value even mean? This has to do with the idea of orientation. For example, imagine your plane as a sheet of paper and you draw your vectors on it with such thick ink that it bled through the paper. If you were to flip this paper and look at the new vectors, this would have a negative determinant. I.e., inverting the "orientation of space"; things that were previously clockwise are now counter-clockwise. This has a similar analog in 3 dimensions: if your basis is right-handed (a la right-hand rule) in one basis, you have switched to a left-handed orientation if your determinant is negative.

Some general laws of determinants are as follows:

- If each element of any row (or column) of a determinant is multiplied by some scalar α, the value of the determinant is also multiplied by a factor of α.

- A non-homogeneous system of linear equations has a unique solution if and only if the determinant of the system's matrix is nonzero.
- The value of a determinant is zero if
  1. all elements of one row (or column) are zero
  2. two rows (or columns) are identical
  3. two rows (or columns) are proportional.
- The determinant of a transpose equals the determinant of the original matrix.
- The determinant of the complex conjugate equals the determinant of the original matrix.
- If two rows (or two columns) of a determinant are switched, the value of the determinant changes sign.
- $\det(\mathrm{AB}) = \det(\mathrm{BA}) = \det(\mathrm{A}) \cdot \det(\mathrm{B})$

2.2 Finding eigenvalues and eigenvectors

Let $\mathbf{A}$ be a linear transformation and $|a\rangle$ a *nonzero* vector such that

$$\mathbf{A}\,|a\rangle = \lambda\,|a\rangle$$

with $\lambda \in \mathbb{C}$. We then say that $|a\rangle$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ. The set of eigenvalues of $\mathbf{A}$ is called the **spectrum** of $\mathbf{A}$. In order to find the eigenvectors and eigenvalues for a given matrix we will rewrite the above equation.

$$(\mathbf{A} - \lambda\mathbf{1})\,|a\rangle = 0$$

If we are to obtain nontrivial solutions to this equation then $(\mathbf{A}-\lambda\mathbf{1})$ must have no inverse. This is only true if:

$$det(\mathbf{A} - \lambda\mathbf{1}) = 0$$

This determinant is a polynomial in λ, called the characteristic polynomial of $\mathbf{A}$. The roots of this polynomial are called the characteristic roots and are the eigenvalues of $\mathbf{A}$. The eigenvectors corresponding the eigenvalues say $\lambda_1, \lambda_2, ..., \lambda_n$. We plug the eigenvalues into above equation:

$$\mathbf{A}\,|a_i\rangle = \lambda_i\,|a_i\rangle \quad \text{where} \quad i \in \{1, 2..n\}$$

You then solve for $|a_i\rangle$. Now for a demonstration. Let us find the eigenvalues and eigenvectors of

$$H = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

We have

$$\det(H - \lambda \mathbf{1}) = \begin{pmatrix} -\lambda & -i \\ i & -\lambda \end{pmatrix} = \lambda^2 - 1$$

Thus the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -1$. To find the eigenvectors we have:

$$0 = (H - \lambda_1 \mathbf{1})\,|a_1\rangle = (H - \mathbf{1})\,|a_1\rangle = \begin{pmatrix} -1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} -\alpha_1 & -i\alpha_2 \\ i\alpha_1 & -1\alpha_2 \end{pmatrix}$$

or $i\alpha_1 = \alpha_2$, which gives $|a_1\rangle = \alpha_1 \begin{pmatrix} 1 \\ i \end{pmatrix}$ where α_1 is an arbitrary complex number. Through the same method we get the second eigenvector $|a_2\rangle = \beta_1 \begin{pmatrix} 1 \\ -i \end{pmatrix}$ where β_1 is an arbitrary complex number. It is usually the case that we want to orthonormalize the eigenvectors. They are already orthogonal so we only need to normalize them.

$$1 = \langle a_1 | a_1 \rangle = \alpha_1^*(1 \ -i)\alpha_1 \begin{pmatrix} 1 \\ i \end{pmatrix} = 2|\alpha_1|^2$$

or $|\alpha_1| = 1/\sqrt{2}$ so this leads to $\alpha_1 = e^{i\phi}/\sqrt{2}$ for some $\phi \in \mathcal{R}$ A similar result is obtained for β_1. The choice $\phi = 0$ leads to

$$|e_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ and } |e_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

2.3 Matrix Inversion

Before you attempt to invert a matrix you should first check to see if it is even possible to do so. Let $\mathbf{A}$ be a square n-by-n matrix. The *Invertible Matrix Theorem* states that all of the following are equivalent; i.e. if any of them are true then all of them are also true.

- $\mathbf{A}$ is invertible, that is, $\mathbf{A}$ has an inverse, is nonsingular, or is nondegenerate.
- $\mathbf{A}$ is row-equivalent to the n-by-n identity matrix $\mathbf{I}_n$.
- $\mathbf{A}$ is column-equivalent to the n-by-n identity matrix $\mathbf{I}_n$.
- $\det \mathbf{A} \neq 0$.
- The equation $\mathbf{A}\mathbf{x} = 0$ has only the trivial solution $\mathbf{x} = 0$.

- The columns of $\mathbf{A}$ are linearly independent.
- The transpose $\mathbf{A}^{\mathrm{T}}$ is an invertible matrix (hence rows of $\mathbf{A}$ are also linearly independent).
- There is an n-by-n matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{I}_n$.
- There is an n-by-n matrix $\mathbf{C}$ such that $\mathbf{CA} = \mathbf{I}_n$.
- The number 0 is not an eigenvalue of $\mathbf{A}$.

This list is even more extensive than the bullet points I've presented here, but exceed the scope of this document.

So let's assume you've satisfied any of these conditions: how do you invert a matrix? Suppose you have some matrix with the following elements:

$$\mathbf{P} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

then its inverse is simply:

$$\mathbf{P}^{-1} = \frac{1}{|\mathbf{P}|} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \tag{2.1}$$

In the case of a 3×3 matrix the process is the same. Take

$$\mathbf{P} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

then its inverse is:

$$\mathbf{P}^{-1} = \frac{1}{|\mathbf{P}|} \begin{pmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{12} \\ a_{33} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{11} \\ a_{32} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix} \tag{2.2}$$

Check for yourself that $\mathbf{P}^{-1}\mathbf{P}$ is indeed the identity matrix. This method can be generalized to some n-by-n matrix as

$$\mathbf{P}^{-1} = \frac{1}{|\mathbf{P}|}\mathbf{C}^{\mathrm{T}} \tag{2.3}$$

Where $\mathbf{C}^{\mathrm{T}}$ is the matrix of cofactors which can be computed via cramer's rule. Some important properties of invertible matrices are

- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- $(k\mathbf{A})^{-1} = k^{-1}\mathbf{A}^{-1}$ for any nonzero scalar k
- $(\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}$
- $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$

2.4 Diagonalization of Matrices

In linear algebra, a square matrix $\mathbf{A}$ is diagonalizable if there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix. There are several reasons one may want a matrix in this form. Aside from its many uses in differential equations, one can calculate the principal axes of inertia, the normal modes of an oscillating system, or easily take many powers of the matrix (indispensable in Markov processes). The importance of these have been elaborated on in section 2.2. Let's look at how to find this invertible matrix $\mathbf{P}$.

We already know that $\mathbf{A}\left|a_i\right\rangle = \lambda_i \left|a_i\right\rangle$; let's explicitly write this out for e.g. a 3×3 matrix:

$$\mathbf{A}[v_1 v_2 v_2] = \mathbf{A}\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \lambda_3 x_3 \\ \lambda_1 y_1 & \lambda_2 y_2 & \lambda_3 y_3 \\ \lambda_1 z_1 & \lambda_2 z_2 & \lambda_3 z_3 \end{pmatrix}$$
$$= \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

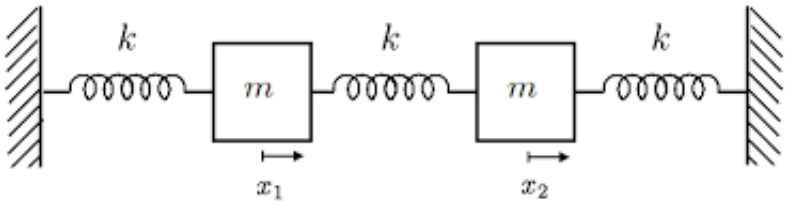
or more compactly:

$$\mathbf{AP} = \mathbf{PD} \tag{2.4}$$

where v_i are the eigenvectors, x, y, z are their components, and $\mathbf{D}$ is the diagonal matrix. All that's left to do is hit both sides from the left with $\mathbf{P}^{-1}$.

2.5 Problems

1. Find the characteristic vibration frequencies for the system of masses and springs shown below.



2. Consider the same system where the middle spring has a spring constant $3k$. Find the characteristic vibration frequencies.

3 Multivariable Calculus

3.1 Partial Derivatives

The basic idea of derivatives and of integrals in two, three, or more dimensions follows the same pattern as for one dimension. They're just more complicated.

The derivative of a function of one variable is defined as

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{3.1}$$

One might then think that the derivative of some multi-dimensional function would be defined as

$$\frac{\partial f(x,y)}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x,y))}{\Delta x} \tag{3.2}$$

and it partially is. Unfortunately, there is a discrepancy with notation here and the idea of taking some $\frac{\partial f}{\partial x}$ is ambiguous about what other variables we're holding constant. I.e. if we were to make a change of coordinates, or go to some three-variable function, the partial derivative itself wouldn't contain enough information to designate exactly what you're varying and what you're holding constant. In some contexts the answer is clear and you won't have any difficulty deciding, but you've already encountered cases for which the distinction is crucial. In thermodynamics, when you add heat to a gas to raise its temperature does this happen at constant pressure or at constant volume or with some other constraint? The specific heat at constant pressure is not the same as the specific heat at constant volume; it is necessarily bigger because during an expansion some of the energy has to go into the work of changing the volume. This sort of derivative depends on type of process that you're using, and for a classical ideal gas the difference between the two molar specific heats obeys the equation

$$c_P - c_v = R$$

If the gas isn't ideal, this equation is replaced by a more complicated and general one, but the same observation applies, that the two derivatives dQ/dT aren't the same.

In thermodynamics there are so many variables in use that there is a standard notation for a partial derivative, indicating exactly which other variables are to be held constant.

$$\left(\frac{\partial U}{\partial V}\right)_T \qquad \text{and} \qquad \left(\frac{\partial U}{\partial V}\right)_P$$

represent the change in the internal energy of an object per change in volume during processes in which respectively the temperature and the pressure are held constant.

3.2 Chain Rule

For functions of one variable, the chain rule allows you to differentiate with respect to still another variable: y a function of x and x a function of t allows

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} \tag{3.3}$$

What happens when you have more variables?

$$\frac{df(x(t),y(t))}{dt} = \left(\frac{\partial f}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial f}{\partial y}\right)_x \frac{dy}{dt} \tag{3.4}$$

As long as you are consistent at abating ambiguity, you can extend this concept to different coordinate systems as well. I.e.

$$\left(\frac{df}{dx'}\right)_{y'} = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial x'}\right)_{y'} + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial x'}\right)_{y'} \tag{3.5}$$

Now, let's see what happens when we switch from cartesian to polar coordinates. The usual form is $x = r\cos\theta$ and $y = r\sin\theta$ so we proceed using Eq. 3.5

$$\begin{aligned}\left(\frac{df}{d\theta}\right)_r &= \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial \theta}\right)_r + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial \theta}\right)_r \\ &= \left(\frac{\partial f}{\partial x}\right)_y (-r\sin\theta) + \left(\frac{\partial f}{\partial y}\right)_x (r\cos\theta)\end{aligned}$$

Now let's say $f(x,y) = x^2+y^2$ so the function is an infinite number of concentric circles. At a fixed radius, what should $df/d\theta$ be? The obvious answer is zero, but let's check with our new equation.

$$\begin{aligned}\left(\frac{df}{d\theta}\right)_r &= \left(\frac{\partial f}{\partial x}\right)_y (-r\sin\theta) + \left(\frac{\partial f}{\partial y}\right)_x (r\cos\theta) \\ &= (2x)(-r\sin\theta) + (2y)(r\cos\theta) \\ &= (-2r^2\cos\theta\sin\theta) + (2r^2\sin\theta\cos\theta) \\ &= 0\end{aligned}$$

3.3 Gradient

The gradient, commonly represented as ∇, is a multi-variable generalization of the derivative. Whereas the ordinary derivative of a function of a single variable is a scalar-valued function, the gradient of a function of several variables is a vector-valued function. Specifically, the gradient of a differentiable function f of several variables, at a point P, is the vector whose components are the partial derivatives of f at P. I.e. At each point in space it provides a vector showing the "steepness" of f at that point and the direction in which f is changing most rapidly.

The electric field can be described in terms of a gradient. For a single point charge at the origin the electric field is

$$\vec{E}(x,y,z) = \frac{kq}{r^2}\hat{r} \tag{3.6}$$

where $\hat{r}$ is the unit vector pointing away from the origin and r is the distance to the origin. This vector can be written as a gradient. Because this $\vec{E}$ is

everywhere pointing away from the origin, it's everywhere perpendicular to the sphere centered at the origin.

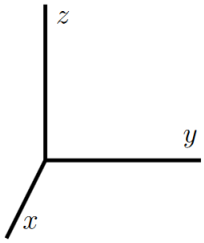
$$\vec{E} = -\nabla \frac{kq}{r}$$

We will see how the gradient is defined in a general coordinate system later, but for now let's define it in Cartesian.

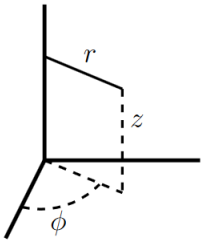
$$\nabla = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z} \tag{3.7}$$

3.4 Cylindrical, Spherical Coordinates

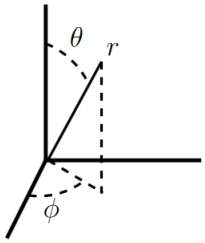
The three common coordinate systems used in three dimensions are rectangular, cylindrical, and spherical coordinates, and these are the ones you have to master.



$-\infty < x < \infty$
$-\infty < y < \infty$
$-\infty < z < \infty$



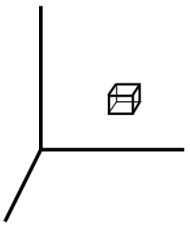
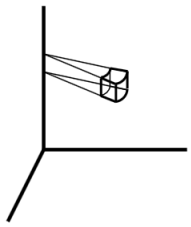
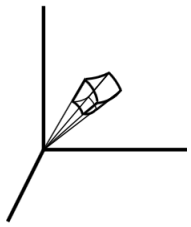
$0 < r < \infty$
$0 < \phi < 2\pi$
$-\infty < z < \infty$



$0 < r < \infty$
$0 < \theta < \pi$
$0 < \phi < 2\pi$

The surfaces that have constant values of these coordinates are planes in rectangular coordinates; planes and cylinders in cylindrical; planes, spheres, and cones in spherical. In every one of these cases the constant-coordinate surfaces intersect each other at right angles, hence the name "orthogonal coordinate" systems. In spherical coordinates I used the coordinate θ as the angle from the z-axis and ϕ as the angle around the axis. In mathematics books these are typically reversed, so watch out for the notation. On the globe of the Earth, ϕ is like the longitude and θ like the latitude.

The volume elements for these systems come straight from the drawings, just as the area elements do in plane coordinates. In every case you can draw six surfaces, bounded by constant coordinates, and surrounding a small box. Because these are orthogonal coordinates you can compute the volume of the box easily as the product of its three edges.

		rectangular	cylindrical	spherical
volume	$d^3r =$	$dx\,dy\,dz$	$r\,dr\,d\phi\,dz$	$r^2\sin\theta\,dr\,d\theta\,d\phi$
area	$d^2r =$	$dx\,dy$	$r\,d\phi\,dz$ or $r\,d\phi\,dr$	$r^2\sin\theta\,d\theta\,d\phi$

3.5 Multiple Integrals

Even in rectangular coordinates integration can be tricky. That's because you have to pay attention to the limits of integration far more closely than you do for simple one dimensional integrals. We'll illustrate this with two dimensional rectangular coordinates first, and will choose a problem that is easy but still shows what you have to look for.

Let's find the area between the curves $y = x$ and $y = x^2/a$. Which do you take the integral of first? It doesn't matter (barring some non-measurable functions, etc.). Let's check both ways.

$$\int_0^a dx \int_{x^2/a}^a dy \stackrel{?}{=} \int_0^a dy \int_y^{\sqrt{ya}} dx$$

On the left hand side, the lower limit of the dy integral comes from the specified equation of the lower curve. The upper limit is the value of y for the given x at the upper curve. After that the limits on the sum over dx comes from the intersection of the two curves: $y = x = x^2/a$ gives $x = a$ for that limit. On the right hand side, I fix y and sum over dx first. The left limit is easy, $x = y$, and the upper limit comes from solving $y = x^2/a$ for x in terms of y. When that integral is done, the remaining dy integral starts at zero and goes up to the intersection at $y = x = a$.

$$\int_0^a dx \int_{x^2/a}^a dy = \int_0^a dx \left[x - \frac{x^2}{a}\right] = \frac{a^2}{2} - \frac{a^3}{3a} = \frac{a^2}{6}$$

$$\int_0^a dy \int_y^{\sqrt{ya}} dx = \int_0^a dy \left[\sqrt{ya} - y\right] = a^{1/2}\frac{a^{3/2}}{3/2} - \frac{a^2}{2} = \frac{a^2}{6}$$

Let's take a physics example. The moment of inertia about an axis is $\int r_\perp^2\,dm$. Here, $r_\perp$ is the perpendicular distance to the axis. What is the moment of inertia of a uniform sheet of mass M in the shape of a right triangle of sides a and b? Take the moment about the right angled vertex. The area mass density,

$\sigma = dm/dA$ is $2M/ab$. The moment of inertia is then

$$\begin{aligned}\int (x^2+y^2)\sigma dA &= \int_0^a dx \int_0^{b(a-x)/a} dt\sigma(x^2+y^2) = \int_0^a dx\sigma \left[x^2 y + y^3/3\right]_0^{b(a-x)/a} \\ &= \int_0^a dx\sigma \left[x^2 \frac{b}{a}(a-x) + \frac{1}{3}\left(\frac{b}{a}\right)^3 (a-x)^3\right] \\ &= \sigma\left[\frac{b}{a}\left(\frac{a^4}{3} - \frac{a^4}{4}\right) + \frac{1}{3}\left(\frac{b^3}{a^3}\frac{a^4}{4}\right)\right] \\ &= \frac{1}{12}\sigma(ba^3 + ab^3) = \frac{M}{6}(a^2+b^2)\end{aligned}$$

Again, the upper bound for the dy integral comes from the equation $y(x) = -(b/a)x + b$. As you build your visual intuition of how to treat multivariable integrals, understanding the limits of integration will become much easier.

3.5.1 The Jacobian

The Jacobian presents an algebraic way of finding the volume elements for more abstract coordinate systems. Suppose we have some triple integral

$$\iiint f(x,y,z)dxdydz$$

Now let's say we want to express the integrand as some function $f(u(x,y,z), v(x,y,z), w(x,y,z))$. Here, we define the Jacobian $\mathcal{J}$ as

$$\mathcal{J} = \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix} \tag{3.8}$$

If we take the determinant of this matrix we can express our triple integral in the new coordinate system as

$$\iiint f(u,v,w)|J| \cdot dudvdw \tag{3.9}$$

3.6 Gradient in other Coordinates

The equation for the gradient computed in rectangular coordinates is Eq. 3.7. How do you compute it in cylindrical or spherical coordinates? You do it the same way that you got 3.5, this time with more coordinates (r, ϕ, and z, etc.). So let's first look at cylindrical

$$df = df(r,\phi,z,dr,d\phi,dz) = \left(\frac{\partial f}{\partial r}\right)_{\phi,z} dr + \left(\frac{\partial f}{\partial \phi}\right)_{r,z} d\phi + \left(\frac{\partial f}{\partial z}\right)_{r,\phi} dz \tag{3.10}$$

We know that in this coordinate system

$$d\vec{r} = \hat{r}dr + \hat{\phi}rd\phi + \hat{z}dz \tag{3.11}$$

The part in the $\hat{\phi}$ direction is the *displacement* of $\mathrm{d}\vec{r}$ in that direction. As ϕ changes by a small amount the distance moved is not $d\phi$; it is $rd\phi$ (visualize geometrically why this is). The equation

$$df = \nabla f \cdot d\vec{r} \tag{3.12}$$

combined with the one above yields

$$\nabla f = \hat{r}\frac{\partial f}{\partial r} + \hat{\phi}\frac{1}{r}\frac{\partial f}{\partial \phi} + \hat{z}\frac{\partial f}{\partial z} \tag{3.13}$$

In spherical coordinates the procedure is identical. All that you have to do is to identify what $d\vec{r}$ is.

$$d\vec{r} = \hat{r}dr + \hat{\theta}rd\theta + \hat{\phi}r\sin\theta d\phi \tag{3.14}$$

Again with this case you have to look at the distance moved when the coordinates changes by a small amount. Just as with cylindrical coordinates this determines the gradient in spherical coordinates.

$$\nabla f = \hat{r}\frac{\partial f}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial f}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi} \tag{3.15}$$

By the end of your core courses you should be able to derive and utilize these equations seamlessly.

3.7 Extrema

With one variable you can look for a maximum or a minimum by taking a derivative and setting it to zero. For several variables you do it several times so that you will get as many equations as you have unknown coordinates. Put this in the language of gradients: $\nabla f = 0$. The derivative of f vanishes in every direction as you move from such a point. As examples,

$$f(x,y) = x^2 + y^2, \quad \text{or} \quad = -x^2 - y^2, \quad \text{or} \quad = x^2 - y^2$$

For all three of these the gradient is zero at $(x, y) = (0, 0)$; the first has a minimum there, the second a maximum, and the third neither – it is a "saddle point." The generic term for all three of these is "critical point."

An important example of finding a minimum is "least square fitting" of functions. How close are two functions to each other? The most commonly used, and in every way the simplest, definition of the distance (squared) between f and g on the interval $a < x < b$ is

$$\int_a^b dx|f(x) - g(x)|^2 \tag{3.16}$$

This means that a large deviation of one function from the other in a small region counts more than smaller deviations spread over a larger domain. The square sees to that.

3.7.1 Hessian

When there's just one variable there is a simple rule that lets you decide. Check the second derivative. If it's positive you have a minimum; if it's negative you have a maximum. If it's zero you have more work to do. Is there a similar method for several variables? Yes, and I'll show it explicitly for two variables. Once you see how to do it in two dimensions, the generalization to N is just a matter of how much work you're willing to do (or how much computer time you can use).

The Taylor series in two variables is (to second order):

$$f(x+dx, y+dy) = f(x,y) + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2 + \cdots \tag{3.17}$$

This can be rewritten as

$$f(\vec{r} + d\vec{r}) = f(\vec{r}) + \nabla f \cdot d\vec{r} + \langle d\vec{r}|H|d\vec{r}\rangle + \cdots$$

The part with the gradient is familiar, and to have either a minimum or a maximum, that will have to be zero. The next term introduces a new idea, the Hessian, constructed from all the second derivative terms. Write these second order terms as a matrix to see what they are, and in order to avoid a lot of clumsy notation use subscripts as an abbreviation for the partial derivatives.

$$\langle d\vec{r}|H|d\vec{r}\rangle = \begin{pmatrix} dx & dy \end{pmatrix} \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \quad \text{where} \quad d\vec{r} = \hat{x}dx + \hat{y}dy \tag{3.18}$$

This matrix is symmetric because of the properties of mixed partials. How do I tell from this whether the function f has a minimum or a maximum (or neither) at a point where the gradient of f is zero? Eq. 3.18 describes a function of two variables even after I've fixed the values of x and y by saying that $\nabla f = 0$. It is a quadratic function of dx and dy . Expressed in the language of vectors this says that f has a minimum if 3.18 is positive no matter what the direction of $\mathrm{d}\vec{r}$ is – H is *positive definite*.

This is a 2×2 symmetric matrix sandwiched inside a scalar product.

$$h(x,y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{3.19}$$

Is h positive definite? That is, positive for all x, y? If this matrix is diagonal it's much easier to see what is happening, so diagonalize it. Find the eigenvectors and use those for a basis.

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{requires} \quad \det \begin{pmatrix} a-\lambda & b \\ b & c-\lambda \end{pmatrix} = 0 \tag{3.20}$$

For the applications here all the a, b, c are the real partial derivatives, so the eigenvalues are real and the only question is whether the λs are positive or negative, because they will be the (diagonal) components of the Hessian matrix

in the new basis. If this is a double root, the matrix was already diagonal. You can verify that the eigenvalues are positive if $a > 0$, $c > 0$, and $4ac > b^2$, and that will indicate a minimum point.

Geometrically the equation $z = h(x, y)$ from Eq. 3.19 defines a surface. If it is positive definite the surface is a paraboloid opening upward. If negative definite it is a paraboloid opening down. The mixed case is a hyperboloid – a saddle.

3.8 Lagrange Multipliers

This is an incredibly clever method to handle problems of maxima and minima in several variables when there are constraints. An example: "What is the largest rectangle?" obviously has no solution, but "What is the largest rectangle contained in an ellipse?" does. I'll describe this procedure for two variables; it's the same for more. The problem stated is that I want to find the maximum (or minimum) of a function $f(x, y)$ given the fact that the coordinates x and y must lie on the curve $\phi(x, y) = 0$. If you can solve the ϕ equation for y in terms of x explicitly, then you can substitute it into f and turn it into a problem in ordinary one variable calculus. What if you can't?

Look at the gradient of f and the gradient of ϕ. The gradient vectors are perpendicular to the curves $f =$ constant and $\phi =$ constant respectively, and at the point where the curves are tangent to each other these gradients are in the same direction (or opposite). Either way one vector is a scalar times the other.

$$\nabla f = \lambda \nabla \phi \tag{3.21}$$

Or more explicitly, the three equations we're solving are

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial \phi}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} - \lambda \frac{\partial \phi}{\partial y} = 0, \qquad \text{and} \qquad \phi(x, y) = 0 \tag{3.22}$$

There are three equations and three unknowns (x, y, λ), and these are the equations to solve for the position of the maximum or minimum value of f.

Let's go back to the original question: What is the largest rectangle that you can inscribe in an ellipse? Let the ellipse and the rectangle be centered at the origin. The upper right corner of the rectangle is at (x, y), then the area of the rectangle is

$$\text{Area} = f(x, y) = 4xy,$$
$$\text{with constraint } \phi(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

So let's proceed with Eq. 3.21. The set of equations we need to solve is

$$4y - \lambda \frac{2x}{a^2} = 0, \qquad 4x - \lambda \frac{2y}{a^2} = 0, \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

The solutions to these three equations are straight-forward. They are $x = a/\sqrt{2}$, $y = b/\sqrt{2}$, $\lambda = 2ab$. The maximum area is then $4xy = 2ab$. The Lagrange multiplier turns out to be the required area.

3.9 Problems

1. Let $w = xy$ and $x = yz$

   (a) Write w purely in terms of x and z, and then purely in terms of y and z.

   (b) Compute the partial derivatives

   $$\left(\frac{\partial w}{\partial x}\right)_y \quad \text{and} \quad \left(\frac{\partial w}{\partial x}\right)_z$$

   (c) Compute the rest of the partial derivatives of w and show that it matters which variable you hold constant.

2. Measured heat capacities of solids and liquids are almost always at constant pressure, not constant volume. To see why, estimate the pressure needed to keep V fixed as T increases, as follows.

   (a) First imagine slightly increasing the temperature of a material at constant pressure. Write the change in volume ,dV_1, in terms of dT and the thermal expansion coefficient $\beta \equiv \frac{\Delta V/V}{\Delta T}$.

   (b) Now imagine slightly compression the material, holding its temperature fixed. Write the change in volume for *this* process, dV_2, in terms of dP and the isothermal compressibility κ_T, defined as

   $$\kappa_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$$

   (c) Finally, imagine that you compress the material just enough in part (b) to offset the expansion in part (a). Then the ratio of dP to dT is equal to $(\partial P/\partial T)_V$, since there is no net change in volume. Express this partial derivative in terms of β and κ_T. Then express it more abstractly in terms of the partial derivatives used to define β and κ_T.

   (d) Compute β, κ_T, and $(\partial P/\partial T)_V$ for an ideal gas and check that the three expressions satisfy your answer in part (c).

   (e) For water at 25° C, $\beta = 2.57\times10^{-4}K^{-1}$ and $\kappa_T = 4.52\times10^{-10}Pa^{-1}$. Suppose you increase the temperature of some volume of water from 20° C to 30° C. How much pressure must you apply to prevent it from expanding? Repeat the calculation for mercury, for which (at 25° C) $\beta = 1.81\times10^{-4}K^{-1}$ and $\kappa_T = 4.04\times10^{-11}Pa^{-1}$. Given the choice, would you rather measure the heat capacities of these substances at constant volume or at constant pressure?

3. Find the potential outside an infinitely long metal pipe, of radius R, placed at right angles to an otherwise uniform electric field E_0. Find the surface charge induced on the pipe.

4. A charge density of $\sigma(\theta) = \rho_0 \sin(5\theta)$ is on the surface of an infinite cylinder of radius R. Find the potential inside and outside the cylinder.

5. Given $s(v,T)$ and $v(\rho,T)$, we define $c_p = T(\partial s/\partial T)_p$, $c_v = T(\partial s/\partial T)_v$ as the specific heat at constant pressure and volume respectively. Show that

$$c_p - c_v = T\left(\frac{\partial s}{\partial v}\right)_T \left(\frac{\partial v}{\partial T}\right)_p$$

6. A roof gutter is to be made from a long strip of sheet metal, 24 cm wide, by bending up equal amounts at each side at equal angles. Find the angle and the dimensions that will maximize the carrying capacity of the gutter.

7. Imagine (draw) a rectangle with two isosceles triangles on its end. What proportions will maximize the area of this shape if the perimeter is fixed? I.e., find the length of the rectangle relative to the sides of the triangle, and the optimal angle the triangle makes with the rectangle.

8. Assume that the Earth is a perfect sphere. Suppose that a rope completes a great circle around the globe such that it hugs the sphere tightly. If we added 2 ft to the rope and held it parallel to the surface at all points, how high above the Earth would the rope be suspended? Would you be able to crawl under it?

9. Calculate the Volume of a sphere in cylindrical coordinates

10. Calculate the electrostatic energy of a charge Q distributed uniformly throughout a sphere of radius R.

11. The surface charge density, dq/dA, on a sphere of radius R is $\sigma(\theta,\phi) = \sigma_0 \sin^2\theta \cos^2\phi$. What is the total charge on the sphere?

12. You have several particles with different allowed energies. If the total number of particles and the total energy are given, how are the particles distributed among the different energies?

    hint: use Stirling's Approximation in conjunction with Lagrange multipliers

$$\ln(n!) \sim \ln\sqrt{2\pi n} + n\ln n - n \tag{3.23}$$

4 Ordinary Differential Equations

4.1 Linear Differential Equations

A differential equation is an expression in which derivatives of an unknown function are set equal to a known function. E.g.

$$\frac{df(x)}{dx} + xf(x) = \cos(x)$$

is a differential equation. To determine the solution of a differential equation we need some initial conditions for the system; i.e. the value of $f(x)$ at some point x to give us a reference point. Every differential equation involves an operator. Operators generally map functions to functions, whereas functions map numbers to numbers. The order of a differential operator is defined by the order of the highest derivative contained in it. Let's look at a specific case of operators. L is a linear operator if and only if

$$L(\alpha f + \beta g) = \alpha L(f) + \beta L(g) \tag{4.1}$$

To surprising accuracy, many physical systems can be expressed by linear differential equations $L(f) = 0$. Maxwell's equations of electrostatics is an obvious application of this technique. The governing equation which generates a Newtonian gravitational field is also a linear differential equation.

The general solution of a differential equation is one that contains a sufficient amount of arbitrary constants to allow it to be a*any* solution of the equation if these constants are assigned appropriate values. As long as the number of initial conditions is at least the number of derivatives, the solution to the differential equation $Lf = 0$ can be uniquely determined.

Given the general solution we can construct n particular solutions $f_1, \ldots, f_n$ as follows: let f_1 be the solution in which the first arbitrary constant, k_1, is unity and the others zero, f_2 be the solution in which the second constant, k_2, is unity and the other zero, etc. By inspection, the general solution is

$$f(x) = \sum_{r=1}^{n} k_r f_r(x) \tag{4.2}$$

Viz., the general solution is a linear combination of n particular solutions (solutions with no arbitrary constant)

Now what do we do if the the stimuli are not encoded in the initial conditions? It is often convenient to formulate a physical problem so that at least some of the stimuli are encoded by a function that we set our differential operator equal to. Such as

$$Lf(x(t)) = g(t) \tag{4.3}$$

Since we are dealing with linear equations, we can distill this into a simpler problem. Let f_1 be the general solution to $Lf = 0$ and f_0 is any solution of $Lf = h$. This f_0 is called the particular integral, while f_1 is called the complementary function. Thus, the general solution to $Lf = h$ is

$$L(f_0 + f_1) = h \tag{4.4}$$

Any first-order linear equation can be written as

$$\dot{f} + qf = h \tag{4.5}$$

The general solution will have one arbitrary constant. It can be found by finding the function $I(x)$ such that

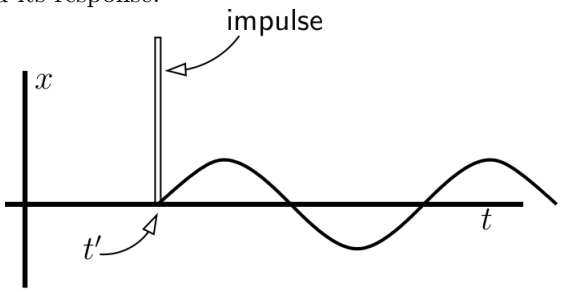
$$I\frac{df}{dx} + Iqf = \frac{dIf}{dx} = Ih \qquad \rightarrow \qquad f(x) = \frac{1}{I(x)} \int_{x_0}^{x} I(x')h(x')dx' \tag{4.6}$$

where x_0 is the required arbitrary constant in the solution, and I is the integrating factor. We need $Iq = dI/dx$, so

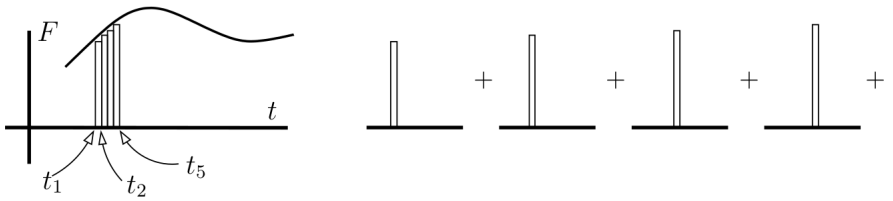
$$\ln I = \int q dx \qquad \rightarrow \qquad I = \exp\left\{\int q dx\right\} \tag{4.7}$$

4.2 Green's Function

The Fourier series was useful in obtaining full solutions for a driving force which is periodic. We will now study an even more powerful tool called the *Green's Function method* which can be used for any driving force. Consider the impulse and its response:



But this impulse by itself is useless; we need to consider more than a single kick to understand the *driving* part of the force. When the external force is the sum of two terms, the total solution is the sum of the solutions for the individual forces. If an impulse at one time gives a solution, an impulse at a later time gives a solution that starts its motion at that later time. The key fact about the equation that you're trying to solve is that it is linear, so you can get the solution for two impulses simply by adding the two simpler solutions. This method can be extended to as many forces as you'd like. E.g.



To complete this idea, the external force is the sum of many terms, the force between t_1 and t_2, that between t_2 and t_3 etc. The total response is the sum of all these individual responses. The idea behind this is that you can think of any force as a sequence of short, small kicks. In fact, because of the atomic nature of matter, that's not so far from the truth. If you can figure out the result of an impact by one molecule, you can add the results of many such kicks to get the answer for N_A molecules.

Let's take some linear operator

$$D(t) = \frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2 \tag{4.8}$$

Where

$$D(t)x(t) = f(t) = \int dt' f(t')\delta(t - t') \tag{4.9}$$

Let's call the response to a delta function impulse a *Green's function*, G, such that

$$D(t)G(t, t') = \delta(t - t') \tag{4.10}$$

Then D becomes

$$D(t)x(t) = \int dt' f(t')D(t)G(t, t') \tag{4.11}$$

Since D is independent of t' we can pull it out of the integral and retrieve an equation for x

$$x(t) = \int dt' f(t')G(t, t') \tag{4.12}$$

That is, as long as we can find a Green's function, we can find a particular solution to any driving force $f(t)$. Moreover, in finding $x(t)$, we are not concerned with initial conditions, we only need to find one form of Green's function, as long as it satisfies Eq. 4.10. As G must satisfy causality, the response function $G = 0$ for $t < t'$ implies that there can be no response before the impulse is given.

Let's take an example for the simple harmonic oscillator. The most basic case assumes no damping so for our linear operator D, $\beta = 0$ and thus

$$D(t) = \frac{d^2}{dt^2} + \omega_0^2$$

For ease of calculation, let's say the impulse occurs at time $t' = 0$. This yields

$$\frac{d^2G(t)}{dt^2} + \omega_0^2 G(t) = \delta(t)$$

Now let's impose some boundary conditions. For $t > 0$ we retrieve the well-known differential equation

$$\ddot{G}_+(t) + \omega_0^2 G(t) = 0$$

which has the general solution

$$G_+(t) = A\cos(\omega_0 t) + B\sin(\omega_0 t) \tag{4.13}$$

First, let's use the causality boundary condition to find

$$G_+(t = 0) = G_-(t = 0) = 0 \quad \text{thus} \quad G_+(t = 0) = A = 0$$

And we're one step closer to the general solution.

$$G_+(t) = B\sin(\omega_0 t)$$

Second, let's use Eq. 4.9 near the boundary $t = 0$

$$\lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} \left(\frac{d^2 G(t)}{dt^2} + \omega_0^2 G(t) \right) dt = \lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} \delta(t) dt$$

This gives us the relation

$$\lim_{\epsilon \to 0} \dot{G}_+(\epsilon) - \dot{G}_-(-\epsilon) + 2\epsilon\omega_0^2 G(0) = 1$$

or

$$B\omega_0 = 1$$

So finally, adding back in the t' in case we don't want to start at time zero,

$$G(t, t') = \begin{cases} \frac{1}{\omega_0} \sin[\omega_0(t - t')] & t > 0 \\ 0 & t \leq 0 \end{cases}$$

Therefore, the solution $x(t)$ for

$$\ddot{x}(t) + \omega_0^2 x(t) = f(t)$$

is given as

$$\begin{aligned} x(t) &= \int_{-\infty}^{\infty} dt' f(t') G(t, t') \\ &= \int_{-\infty}^{t} dt' f(t') G(t, t') \end{aligned}$$

since $G(t)$ is only non-zero when $t > t'$

$$x(t) = \frac{1}{\omega_0} \int_{-\infty}^{t} dt' f(t') \sin[\omega_0(t - t')] \tag{4.14}$$

As you can guess, Green's function is a very powerful tool you'll encounter in many areas of physics.

4.3 Separation of Variables

If you have a first order differential equation $Lf(x(t), t) = 0$ and can decompose this into an equation of the form $g(x)dx = h(t)dt$ then you have successfully "separated variables." Now all you have to do is integrate.

For example, the total energy in an undamped harmonic oscillator is $E = mv^2/2 + kx^2/2$. Solve for dx/dt and you get

$$\frac{dx}{dt} = \sqrt{\frac{2}{m}(E - kx^2/2)}$$

So the obvious form we want is

$$\int \frac{dx}{\sqrt{\frac{2}{m}(E - kx^2/2)}} = \int dt$$

Using some substitution we arrive at the familiar result for a simple harmonic oscillator

$$x(t) = \sqrt{\frac{2E}{k}} \sin\left(\sqrt{\frac{k}{m}}t + \phi\right) \tag{4.15}$$

This is a very simple example so we will elaborate a little more in the PDEs section.

4.4 Systems of Linear Differential Equations

Many physical systems require more than one variable to quantify their configuration: for example a circuit might have two connected current loops, so one needs to know what current is flowing in each loop at each moment. A set of differential equations is necessary for each loop. The process of solving these equations is very similar to that of the single linear differential equation. To reiterate:

1. Separate the functions of the unknown variable on one side and the known function on the other.

2. Find the general solution (complementary function) in the same way as in Eq. 4.6.

3. Find a particular integral by making an ansatz (trial solution) for each term. Experiment with different candidates; many times it is a polynomial or exponential (or some combination of them). Other times, the solution to a differential equation is so common in nature it gets its own name (Bessel, Legendre, Spherical Harmonics, etc.)!

4.5 Laplace Transform

Laplace transforms are second only to Fourier transforms in finding solutions to physical problems. The Laplace transform of a function $f(t)$ is denoted

$$\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t) dt \tag{4.16}$$

You'll notice that the Laplace transform (LT) of a function is a complex function of a complex variable, while the Fourier transform of a function is a complex function of a real variable. By now, you should know how to tackle several integrals of this form, but for those that aren't obvious, the LT has many known solutions to common functions (e.g. Bessel) and can be looked up in a table.

More importantly, how does the LT help us solve differential equations? Suppose that $f(t)$ and all its derivatives are continuous. Then,

$$\mathcal{L}\{f^{(n)}\} = s^n F(s) - s^{n-1} f(0) - s^{n-2} f'(0) - \cdots - f^{(n-1)}(0) \tag{4.17}$$

Where $F(s) = \mathcal{L}\{f(t)\}$. One can immediately see the utility of this technique for initial value problems. Let's solve an example initial value problem:

$$y'' - 10y' + 9y = 5t, \quad y(0) = -1, \quad y'(0) = 2$$

The process here is to take a LT of both sides, algebraically manipulate the equations into a more tractable form, then take the inverse LT to find the solution.

$$\begin{aligned} &\mathcal{L}\{y''\} - 10\mathcal{L}\{y'\} + 9\mathcal{L}\{y\} = 5\mathcal{L}\{t\} \quad \rightarrow \\ &s^2 Y(s) - sy(0) - y'(0) - 10(sY(s) - y(0)) + 9Y(s) = \frac{5}{s^2} \end{aligned}$$

Where $Y(s)$ is simply $\mathcal{L}\{y\}$. Plugging in the initial conditions yields a solution for $Y(s)$:

$$Y(s) = \frac{5}{s^2(s-9)(s-1)} + \frac{12-s}{(s-9)(s-1)} = \frac{5 + 12s^2 - s^3}{s^2(s-9)(s-1)}$$

This can be partial fraction decomposed into some form as

$$Y(s) = \frac{A}{s} + \frac{B}{s^2} + \frac{C}{s-9} + \frac{D}{s-1}$$

Some algebra later (i.e. matching the numerators for both sides) and we find

$$Y(s) = \frac{50}{81s} + \frac{5}{9s^2} + \frac{31}{81(s-9)} - \frac{2}{s-1}$$

So all that's left is to take the inverse LT. Doing this we get

$$y(t) = \frac{50}{81} + \frac{5}{9}t + \frac{31}{81}e^{9t} - 2e^t$$

Clearly the LT is invaluable in solving all kinds of differential equations; I challenge anyone to develop this ansantz by eye.

4.6 Problems

1. Find the distance which an object moves in time t if it starts from rest and has an acceleration $d^2x/dt^2 = ge^{-kt}$. Look at the behavior of acceleration for very small and very large t. What does it look like? Hint: this equation roughly models the acceleration on a parachutist.

2. Find the position x of a particle as a function of time t if its acceleration is $d^2x/dt^2 = A\sin\omega t$

3. The momentum p of an electron at velocity v near the velocity c of light increases according to the formula

$$p = \frac{m_0 v}{\sqrt{1 - v^2/c^2}}$$

where m_0 is the constant rest mass. If an electron is subject to a constant force F, Newton's second law describing its motion is

$$\frac{dp}{dt} = \frac{d}{dt}\left(\frac{m_0 v}{\sqrt{1 - v^2/c^2}}\right) = F$$

Find the velocity as a function of time and show that the limiting velocity as t tends to infinity is c. Find the distance traveled by the electron in time t if it starts from rest.

4. Suppose the rate at which bacteria in a culture grow is proportional to the number present at any time. Write and solve the differential equation for the number N of bacteria as a function of time t if there are N_0 bacteria when $t = 0$

5. Heat is escaping at a constant rate dQ/dt through the walls of a long, cylindrical pipe. Find the temperature T at a distance r from the axis of the cylinder if the inside wall has a radius $r = 1$ and temperature $T = 100$ and the outside wall has $r = 2$ and $T = 0$.

6. Water with a small salt content (5 lbs/1000 gal) is flowing into a very salty lake at the rate of 10^5 gallons of water per hour. If at some time $t = 0$ the volume of the lake is 10^9 gallons, and its salt content is 10^7 lbs, find the salt content at time t. Assume a homogeneous mixture of salt and that the diffusion timescale is arbitrarily fast.

7. Maybe a rocket problem..? A Rocket of mass $m(t)$ is propelled by steadily ejecting part of its mass at velocity u in the frame of the rocket. Neglecting gravity, the differential equation of the rocket is $m(dv/dm) = -u$ as long as $v \ll$ the speed of light. Find v as a function of m if $m(v = 0) = m_0$.

8. The differential equation for the path of an orbiting planet is

$$\frac{1}{r^2}\frac{d}{d\theta}\left(\frac{1}{r^2}\frac{dr}{d\theta}\right) - \frac{1}{r^3} = -\frac{k}{r^2}$$

Using a change of variables, solve the equation to show that the path is a conic section.

9. catenary problem?

10. Solve the following IVP

$$2y'' + 3y' - 2y = te^{-2t}, \qquad y(0) = -1, \qquad y'(0) = -4$$

5 Vector Calculus

5.1 Index Notation

Index notation is a very useful way of representing operators, vectors, and tensors. It is also makes deriving vector identities very easy and seamless. It also connects nicely to general and special relativity. Where one uses non orthogonal curvilinear coordinates. Familiar cartesian operators in index notation form are

$$\nabla \cdot A_i = \partial_i A_i \qquad \text{Gradient}$$

$$(\nabla \times A)_i = \epsilon_{ijk} \partial_j A_k \qquad \text{Cross product}$$

$$\nabla_i \cdot \nabla_i = \partial_i \partial_i \qquad \text{Laplacian}$$

Here we are using Einstein notation where repeated indices are summed over. Also ϵ_{ijk} is the levi-civita symbol. Properties of the levi-civita symbol are $\epsilon_{ijk} = 1$ if i, j, k is an even permutation of 1,2,3, $\epsilon_{ijk} = -1$ if i, j, k is an odd permutation of 1,2,3, and $\epsilon_{ijk} = 0$ if there are any repeated indices. Levi-Civita symbol is related to the Kronecker delta.

$$\epsilon_{ijk}\epsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}$$

$$= \delta_{il}(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) - \delta_{im}(\delta_{jl}\delta_{kn} - \delta_{jn}\delta_{kl}) + \delta_{in}(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})$$

A special case of the above equation is

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$$

5.2 Identities of Vector Operators

Now using index notation we can start deriving identities of vector operators.

$$\begin{aligned} (\nabla \times (\nabla \times V))_i &= \epsilon_{ijk}\partial_j(\epsilon_{klm}\partial_l v_m) \\ &= \epsilon_{ijk}\epsilon_{klm}\partial_j\partial_l V_m \\ &= \epsilon_{kij}\epsilon_{klm}\partial_j\partial_l V_m \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\partial_j\partial_l V_m \\ &= \partial_i\partial_j V_j - \partial_j\partial_j V_i \\ &= (\nabla(\nabla \cdot V) - \nabla^2 V)_i \end{aligned}$$

In a few lines using index notation we have derived vector identity which would have taken a lot longer otherwise. Please prove the following vector identities

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = (\nabla \times \mathbf{A}) \cdot \mathbf{B} - \mathbf{A} \cdot (\nabla \times \mathbf{B})$$

5.3 Fields

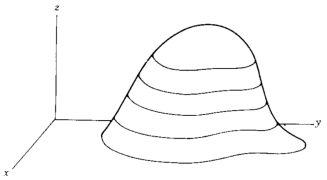
You are probably very familiar with fields already. They are all around us after all. Examples of fields are the electric field around a point charge and the magnetic field induced by a moving charge. Fields appear again in the study of quantum field theory which is a subject that has applications in all areas of physics. The term field is used to refer to both the region and the value of the physical quantity in the region.

Scalar fields, $\phi(x)$, are fields whose values are given by a single value at each point in space-time. An example of a scalar field is temperature. This value does not change under transformations of space. A vector field,$\phi^\mu(x)$, are fields where we attach a vector at each point in space-time. Components of this vector transform themselves contravariantly under rotations in space.

$$\begin{aligned} \phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x) \qquad & \text{scalar fields} \\ \phi^\mu(x) \to \Lambda^\mu_\nu \phi^\nu(\Lambda^{-1}x) \qquad & \text{vector fields} \end{aligned}$$

where Λ^μ_ν is a 4x4 matrix.
We represent an arbitrary Lorentz transformation as

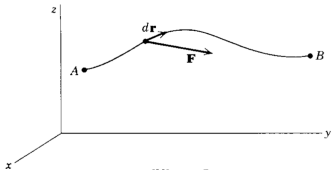
$$x^\mu \to x^{'\mu} = \Lambda^\mu_\nu x^\nu$$



In figure above we have a scalar field. The series of curves on the hill are called contour lines.

5.4 Problems

5.5 Line Integrals

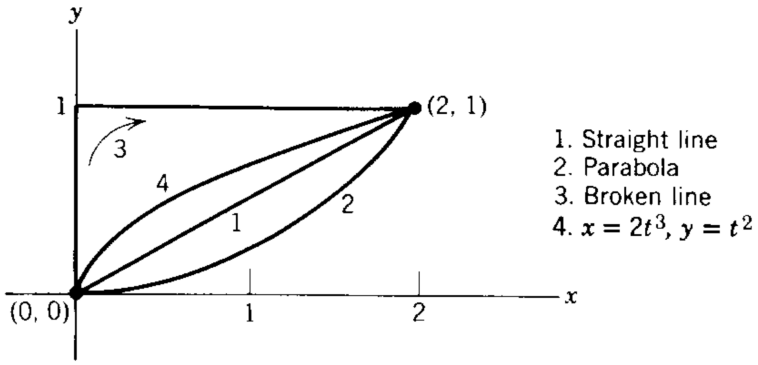


Work done by a force $\mathbf{F}$ on an object which undergoes infinitesimal vector displacement $d\mathbf{r}$ is $dW = \mathbf{F} \cdot d\mathbf{r}$. Let the object move from point A to point B while $\mathbf{F}$ varies as it moves. $\mathbf{F}$ can be a function of x, y, z, however on a curve x, y, z are related by equations of a curve and therefore there is only one independent variable. The integral of $d\mathbf{W} = \mathbf{F} \cdot d\mathbf{r}$ becomes an ordinary integral

of a function of one variable.
A line integral means an integral along a curve that is a single integral.

5.6 Example

Let $\mathbf{F} = xy\hat{i} - y^2\hat{j}$. Find the work done along the 4 paths shown in the figure below. Each path has a different amount of work associated with it. Which path do you think requires the least amount of work? Let's calculate and find out!



The infinitesimal vector displacement is $d\mathbf{r} = \hat{i}dx + \hat{j}dy$ and the work done is $W = \int xydx - y^2dy$

5.7 Path 1: Straight line

We can write y as a function of x. Let y = 1/2 x, and now $dy = 1/2dx$. Now we have one independent variable, x.

$$
\begin{aligned}
W &= \int_0^2 (1/2x^2dx - 1/4x^2(1/2)dx) \\
&= \int_0^2 (1/2 - 1/8)x^2dx \\
&= 3/8 \int_0^2 x^2dx \\
&= 1
\end{aligned}
$$

5.8 Path 2: Parabola

Path 2 is a parabola and the equation can be written as $y = 1/4x^2$ so $dy = \frac{1}{2}xdx$

$$\begin{aligned} W &= \int_0^2 (\frac{1}{4}x^3 - (\frac{1}{4}x^2)^2 \frac{1}{2}x)dx \\ &= \int_0^2 (\frac{1}{4}x^3 - \frac{1}{32}x^5)dx \\ &= \left(\frac{1}{16}x^4 - \frac{1}{192}x^6\right)_0^2 = \frac{2}{3} \end{aligned}$$

Do paths 3, and 4 as an exercise.

5.9 Gauss's Law

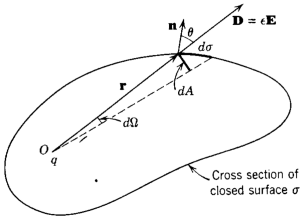
Gauss's law is an application of the divergence theorem. Let's derive Gauss's law from Coulomb's law. Coulomb's law gives for the electric field of a point charge the following.

$$\mathbf{E} = \frac{q}{4\pi\epsilon r^2} e_{\mathbf{r}}$$

where ϵ is the dielectric constant and $k = \frac{1}{4\pi\epsilon}$ is Coulomb's constant which is equal to $9 * 10^9 \frac{N \cdot m^2}{C^2}$.

$$\oint_{\text{closed surface}} D \cdot \sigma = \int_{\text{volume}} \rho d\tau \qquad \text{Gauss's law}$$

The electric displacement is defined by $\mathbf{D} = \epsilon\mathbf{E}$



In the figure above σ is a closed surface surrounding the point charge q. $d\sigma$ is an element of area of the surface at the point $\mathbf{r}$ and $\mathbf{n}$ is the unit normal to $d\sigma$. Let dA be the projection of $d\sigma$ onto a sphere of radius r and center O. Let $d\Omega$

be the solid angle subtended by $d\sigma$ at O
Hence by definition:

$$d\Omega = \frac{1}{r^2} dA$$

and so we have

$$\mathbf{D} \cdot \mathbf{n} d\sigma = D\cos\theta d\sigma = DdA = \frac{q}{4\pi r^2} \cdot r^2 d\Omega = \frac{1}{4\pi} q d\Omega$$

Now integrating this expression over a closed surface σ we have

$$\begin{aligned} \oint_{\text{closed surface } \sigma} D \cdot \mathbf{n} \, d\sigma &= \frac{q}{4\pi} \int_{\text{total solid angle}} d\Omega \\ &= \frac{q}{4\pi} 4\pi \\ &= q \end{aligned}$$

This is gauss's law for a single point charge inside the closed surface σ. For several charges q_i we have the following.

$$\oint_{\text{closed surface } \sigma} D \cdot \mathbf{n} \, d\sigma = \sum_i \oint_{\text{closed surface } \sigma} D_i \cdot \mathbf{n} \, d\sigma = \sum_i q_i$$

For any charge distribution inside a closed surface we have

$$\oint_{\text{closed surface } \sigma} D \cdot \mathbf{n} \, d\sigma = \text{ total charge inside the closed surface} \qquad \text{(Gauss's law)}$$

Now instead if we had a charge distribution with charge density ρ gauss's law would be:

$$\oint_{\text{closed surface } \sigma} D \cdot \mathbf{n} \, d\sigma = \int_{\text{volume bounded by } \sigma} \rho \cdot d\tau \qquad \text{(Gauss's law)}$$

Applying the divergence theorem we have

$$\oint_{\text{closed surface } \sigma} D \cdot \mathbf{n} \, d\sigma = \int_{\text{volume bounded by } \sigma} \nabla \cdot \mathbf{D} \, d\tau \qquad \text{(Divergence theorem)}$$

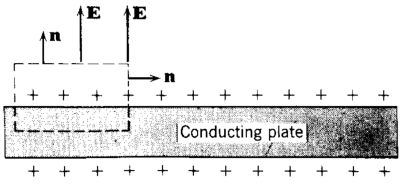
and so

$$\int \nabla \cdot \mathbf{D} \, d\tau = \int \rho \, d\tau$$

Since this is true for any volume τ we have, $\nabla \cdot \mathbf{D} = \rho$. This is one of Maxwell's equations.

5.10 Example

Let's find the electric field $\mathbf{E}$ above a very large conducting plate carrying a surface charge σ. The electric field is zero inside a conductor for an electrostatic problem. By symmetry the $\mathbf{E}$ field must be vertical. We now find $\oint \mathbf{D} \cdot \mathbf{n} d\sigma$ over the box whose cross section is the dotted lines. The integral over the bottom surface is zero since $\mathbf{D} = 0$ inside the conductor. The integral over the vertical sides is zero since $\mathbf{D}$ is perpindicular to it. On the top surface we have $\int \mathbf{D} \cdot \mathbf{n} d\sigma = |\mathbf{D}| \cdot$ surface area. Using Gauss's law we have $|\mathbf{D}| \cdot$ surface area = charge enclosed by the box or $|\mathbf{D}| \cdot$ surface area $= \sigma \cdot$ surface area. Therefore we have $|\mathbf{D}| = \sigma$

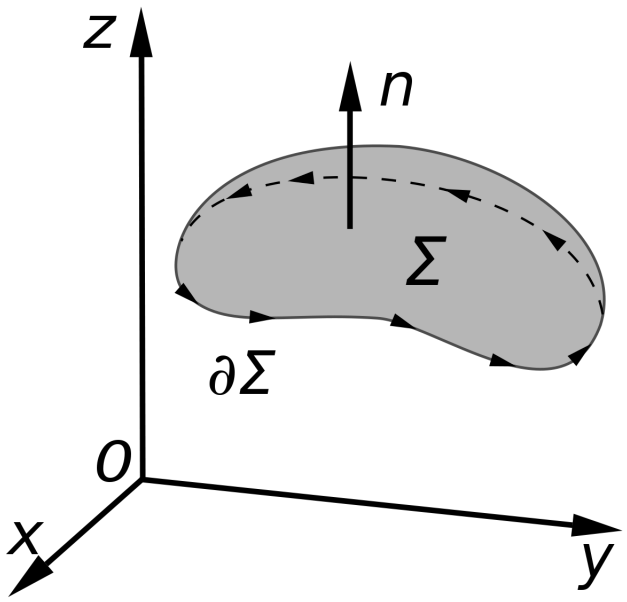


5.11 Problems

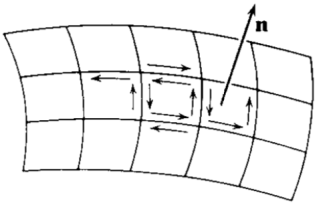
Obtain Coulomb's law from Gauss' law by considering a spherical surface σ with center at q.

5.12 Curl and Stoke's Theorem

Stoke's theorem relates an integral over an open surface to the line integral over a curve bounding a surface. A butterfly net is an example of the above, the net is the surface and the supporting rim is the curve bounding the surface. The surfaces we consider are surfaces we can obtain by deforming a hemisphere. The surfaces are two sided. Stoke's theorem does not work on surfaces that are one sided like a mobius strip.



Consider dividing the surface above into area elements $d\sigma$ by a network of curves as shown below. Draw a unit vector $\mathbf{n}$ perpendicular to each area element.



Now we have using

$$\iint_A (\nabla \times \mathbf{V}) \cdot \mathbf{k}$$

$$\oint_{\text{around } d\sigma} \mathbf{V} \cdot d\mathbf{r} = \iint_{d\sigma} (\nabla \times \mathbf{V}) \cdot \mathbf{n} \; d\sigma$$

for each element. Now if we add all such equations we get

$$\sum_{\text{all}} \oint_{\text{around } d\sigma} \mathbf{V} \cdot d\mathbf{r} = \iint_{d\sigma} (\nabla \times \mathbf{V}) \cdot \mathbf{n} \; d\sigma$$

From the above figure we see that all interior line integrals cancel out because along the border between two $d\sigma$'s the two line integrals cancel out.

$$\oint_{\text{curve bounding } \sigma} \mathbf{V} \cdot d\mathbf{r} = \iint_{d\sigma} (\nabla \times \mathbf{V}) \cdot \mathbf{n} \; d\sigma \qquad \text{Stoke's Theorem}$$

6 Fourier Analysis

Fourier analysis show up throughout all of physics since many problems involving oscillations are common.

6.1 Fourier Series

Lets focus on the elementary periodic functions e^{inx}. We can define the set $\{|e_n\rangle\}_{n=1}^{\infty}$ so that "x components" are given by

$$\langle x|e_n\rangle = \frac{1}{\sqrt{2\pi}} e^{inx}$$

These functions or ket vectors are orthonormal and complete on $\mathcal{L}^2(-\pi, \pi)$. Any piecewise continuous function of x can be expressed as a linear combination of these orthonormal vectors. If $|f\rangle \in \mathcal{L}^2(-\pi, \pi)$, then

$$|f\rangle = \sum_{n=-\infty}^{\infty} f_n |e_n\rangle \text{ where } f_n = \langle e_n | f\rangle$$

We can write

$$\begin{aligned} \langle x|f\rangle &= f(x) \\ &= \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} f_n e^{inx} \end{aligned}$$

The terms f_n are the coefficients and they are given by:

$$\begin{aligned} f_n &= \langle e_n | \, 1 \, | f\rangle \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \end{aligned}$$

The function e^{inx} all have period 2π. If we are given $f(x)$ over any interval of length 2π we can sketch it for that interval and repeat it periodically over with period 2π. However many problems in physics don't come in intervals of 2π. We can define the fourier series on an interval $\{a, b\}$ of length $L = b - a$.

$$\begin{aligned} F(x) &= \frac{1}{\sqrt{L}} \sum_{n=-\infty}^{\infty} F_n e^{2n\pi ix/L} \\ F_n &= \frac{1}{\sqrt{L}} \int_b^a e^{-i(2\pi n/L)x} F(x) dx \end{aligned}$$

6.2 Fourier Series in Higher Dimensions

Fourier series generalize straightforwardly to higher dimensions

$$F(r) = \frac{1}{\sqrt{V}} \sum_k F_k e^{i\mathbf{g_k}\cdot\mathbf{r}}$$

where we used the following notations:

$$\begin{aligned} k &\equiv (k_1, k_2, \ldots, k_N) \\ g_k &= 2\pi(k_1/L, \ldots, k_N/L) \\ V &= L_1 L_2 \cdots L_N \\ F_k &\equiv f_{k_1} \cdots f_{k_N} \\ r &= (x_1, x_2, \cdots x_N) \end{aligned}$$

In one dimension, the shape of the smallest region of periodicity is simply a line segment of length L. In two and more dimensions however smallest region is degenerate and may have different shapes. In two dimensions there can be rectangles, hexagons, pentagons and so forth. Thus we let V stand for a primitive cell of the N-dimensional lattice. This cell in three-dimensions is called the Wigner-Seitz cell. Therefore we have:

$$\begin{aligned} F(r) &= \sum_{\mathbf{k}} F_k e^{i\mathbf{g}_k\cdot\mathbf{r}} \\ F_{\mathbf{k}} &= \frac{1}{V}\int_V F(\mathbf{r}) e^{-i\mathbf{g_k}\cdot\mathbf{r}} d^N x \end{aligned}$$

Where the sum is a multiple sum over $(k_1, k_2, \ldots, k_N)$ and the integral is a multiple integral over the Wigner-Seitz cell.

6.3 Fourier Transform

The fourier transform is defined as

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dx$$

This transform the function from real space to fourier space. The inverse fourier transform is defined as

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx$$

It is useful and straightforward to generalize the fourier transform to higher dimensions

$$f(\mathbf{r}) = \frac{1}{(2\pi)^{n/2}} \int d^n k e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k}),$$
$$\tilde{f}(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}} \int d^n x e^{-i\mathbf{k}\cdot\mathbf{r}} f(\mathbf{r}).$$

Now let us use Dirac notation to illuminate the preceding results

$$\begin{aligned}
\langle \mathbf{r}|\mathbf{k}\rangle &= \frac{1}{(2\pi)^{n/2}} e^{i\mathbf{k}\cdot\mathbf{r}}, \\
\langle \mathbf{k}|\mathbf{r}\rangle &= \frac{1}{(2\pi)^{n/2}} e^{-i\mathbf{k}\cdot\mathbf{r}}, \\
\mathbf{1} &= \int_{-\infty}^{\infty} |k\rangle \langle k| \, dk, \\
\langle \mathbf{k}|\mathbf{k}'\rangle &= \delta(\mathbf{k}-\mathbf{k}') \\
&= \frac{1}{(2\pi)^{n/2}} \int d^n x e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}
\end{aligned}$$

The fourier transform is very useful in solving differential equations. This is because the derivative operator is $\mathbf{r}$ space turns into multiplication in $\mathbf{k}$ space. The n-dimensional gradient is:

$$\nabla f(\mathbf{r}) = \frac{1}{(2\pi)^{n/2}} \int d^n k (i\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k})$$

and the n-dimensional Laplacian is:

$$\nabla^2 f(\mathbf{r}) = \frac{1}{(2\pi)^{n/2}} \int d^n k (-\mathbf{k}^2) e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{f}(\mathbf{k})$$

6.4 Problems

1. Show that for real $f(x)$, $f_{-n} = f_n^*$
2. Find the Fourier tranform of

$$f(x) = \begin{cases} b & \text{if } |x| < a, \\ 0 & \text{if } |x| > a \end{cases}$$

7 Partial Differential Equations

As is typical with differential equations, it is sometimes easier to describe how the setup *changes* from moment to moment than it is to jump straight to a description of the full evolution. This concept is especially useful when quantities

we want to model are functions of other quantities we also want to solve for. Brownian motion, heat flow, or even the evolution of a certain stock's value all use partial differential equations.

How are PDEs different from ODEs? The essential difference is that PDEs require derivatives in multiple variables. Take the simple system of orbiting planets in a plane; the only variable which the system depends on is time so we can typically solve this in the ordinary way. With PDEs, we typically have infinitely many values changing from a continuum. The way that any one of these values changes depends on the other values in the system.

Some common PDEs in physics are Laplace's equation

$$\nabla^2 \phi = 0 \tag{7.1}$$

which is satisfied by the temperature $\phi = \phi(x, y, z)$ in a solid that is in thermal equilibrium, or an electrostatic potential $\phi(x, y, z)$ in a space with no electric charge, etc. An extension of this equation is the general heat equation

$$\frac{\partial \phi}{\partial t} = \alpha \nabla^2 \phi \tag{7.2}$$

which is satisfied by the temperature $\phi = \phi(x, y, z, t)$ of an object which conducts heat, where α is some parameter which dictates the physical properties of the object (e.g. conductivity, emissivity, etc.). The wave equation

$$\frac{\partial^2 \phi}{\partial t^2} = c^2 \nabla^2 \phi \tag{7.3}$$

which models several different phenomenon: vibrations of strings, thin membranes, acoustic waves, etc. The constant c typically represents the speed of propagation of the wave.

7.1 Heat Equation

Before we dive into solving PDEs, let's set up some context to understand a little more physics first. The flow of heat in one dimension is described by the heat conduction equation

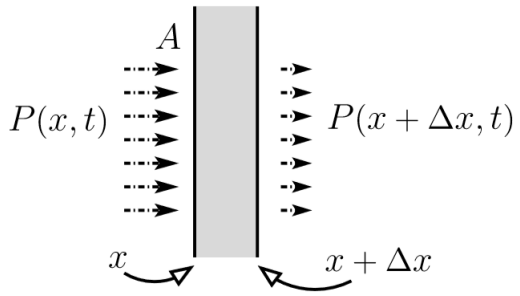
$$P = -\kappa A \frac{\partial T}{\partial x} \tag{7.4}$$

where P is the power in the form of heat energy flowing in a direction x through a wall of area A, and κ is the wall's thermal conductivity. The minus sign indicates that the heat flows from hotter regions to colder regions. To get an intuition for thermal conductivity, imagine you enter your car on a hot summer day. When you sit down and take off the parking brake, your hand is able to put it down before you feel any noticeable effect; but when you barely touch the metal part of your seatbelt, you immediately feel the burning on your skin. These two objects are obviously in equilibrium with the rest of the car, but the metal has a much higher thermal conductivity than the hand-brake. Hence, the heat is transferred much faster.

When more heat comes into a region than leaves it, the temperature will increase. This is captured in the specific heat c.

$$dQ = mcdT, \qquad \text{or} \qquad \frac{dQ}{dt} = mc\frac{dT}{dt} \tag{7.5}$$

Or, succinctly, the temperature rises in a way that's proportional to the heat, and inversely proportional to the mass.



For a slab of area A, thickness δx, and mass density ρ, let the coordinates of the two sides be x and $x + \Delta x$.

$$m = \rho A \Delta x, \qquad \text{and} \qquad \frac{dQ}{dt} = P(x,t) - P(x+\Delta x, t)$$

The net power into this volume is the power in from one side minus the power out from the other. Put these three equations together and you find

$$\frac{dQ}{dt} = mc\frac{dT}{dt} = \rho A \Delta x c \frac{dT}{dt} = \kappa A \frac{\partial T(x+\Delta x, t)}{\partial x} - \kappa A \frac{\partial T(x,t)}{\partial x}$$

Looking at the last equality and dividing by Δx, the form on the right hand side should look familiar to you. We obtain

$$\frac{\partial T}{\partial t} = \frac{\kappa}{\rho c}\frac{\partial}{\partial x}\left(\frac{T(x+\Delta x, t) - T(x,t)}{\Delta x}\right) \overset{\Delta x \to 0}{=} \frac{\kappa}{\rho c}\frac{\partial^2 T}{\partial x^2}$$

Some assumptions were obviously made here, but it's hard to not be cavalier with PDEs. The three-dimensional analog is found in a similar way, but the final equation is

$$\frac{\partial T}{\partial t} = \frac{\kappa}{\rho c}\nabla^2 T \tag{7.6}$$

You can think of the Laplacian as "how different is a point from the average of its neighbors?" So the temperature changes (in time) as a function of the second derivative of how much hotter or colder its neighbors is!

7.2 Separation of Variables

Recall section 4.3. The trick is to start by looking for a solution to the equation in the form of a product of a function of x and a function of t. $T(x,t) = f(t)g(x)$. Do not assume that every solution to the equation will look like this. What will happen is that we'll be able to express every solution as a sum of such factored forms. Let's go back to the one-dimensional heat equation and see if we can develop some solution for the initial temperature distribution.

$$\frac{\partial T}{\partial t} = \frac{\kappa}{\rho c}\frac{\partial^2 T}{\partial x^2} \quad \text{becomes} \quad g\frac{df}{dt} = \frac{\kappa}{\rho c} f\frac{d^2 g}{dx^2}$$

Let C be some constant $C = \kappa/\rho c$ and move the functions to their respective sides to get

$$\frac{1}{f}\frac{df}{dt} = C\frac{1}{g}\frac{d^2 g}{dx^2}$$

The left side of this equation is a function of t alone, no x. The right side is a function of x alone with no t, hence the name separation of variables. Because x and t can vary quite independently of each other, the only way that this can happen is if the two side are constant (the same constant).

$$\frac{1}{f}\frac{df}{dt} = \alpha = C\frac{1}{g}\frac{d^2 g}{dx^2}$$

At this point, the constant α can be anything, even complex. The PDE itself only represents one of the three constraints which our solution must satisfy if it is to accurately describe the system. Besides the PDE, it must also satisfy the boundary conditions at the edges of the system, as well as the initial condition. For a particular specified problem there will be boundary conditions placed on the functions, and those will constrain the αs. If α is real and positive then

$$g(x) = A\sinh\sqrt{\alpha/C}x + B\cosh\sqrt{\alpha/C}x \quad \text{and} \quad f(t) = e^{\alpha t}$$

For a negative, real α, the hyperbolic functions become circular functions.

$$g(x) = A\sin\sqrt{\alpha/C}x + B\cos\sqrt{\alpha/C}x \quad \text{and} \quad f(t) = e^{\alpha t}$$

If $\alpha = 0$ then

$$g(x) = Ax + B \quad \text{and} \quad f(t) = \text{constant}$$

For imaginary α the $f(t)$ is oscillating and the $g(x)$ has both exponential and oscillatory behavior in space. This can (and often does) happen in very ordinary physical situations. This analysis provides a solution to the one-dimensional heat equation, valid for any α. A sum of such solutions for different αs is also a solution, for example

$$T(x,t) = A_1 e^{\alpha_1 t}\sin\sqrt{-\alpha_1/C}x + A_2 e^{\alpha_2 t}\sin\sqrt{-\alpha_2/C}x \tag{7.7}$$

or any other linear combination with various αs

$$T(x,t) = \sum_{\{\alpha\}} f_\alpha(t) g_\alpha(x) \tag{7.8}$$

It is the combined product that forms a solution to the original partial differential equation, not the separate factors. Determining the details of the sum can be found via Fourier series.

7.3 Wave Equation

For simplicity, let's inspect the 1D wave equation between some interval $[0, L]$ where we have some boundary conditions at the ends of the interval. The canonical example is a vibrating string, where the ends are fixed. I.e. $\phi(0,t) = 0$ and $\phi(L,t) = 0$. The problem must also come equipped with some initial conditions of the position and velocity of the string, so $\phi(x,0) = f(x)$ and $\frac{\partial \phi(x,0)}{\partial t} = g(x)$.

We can solve this by separation of variables as before. A solution to

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

has solutions

$$\phi(x,t) = \sum_{n=1}^{\infty} (a_n \cos(n\pi ct/L) + b_n \sin(n\pi ct/L)) \sin(n\pi x/L)$$

where I've implicitly imposed the BCs, forcing the position dependent cosine term to vanish and the poles of the position dependent sine to be integer multiples of π/L. All that's left to do is impose the given initial conditions of the position and velocity to determine a_n and b_n and the problem is solved.

7.4 More on Boundary Conditions

There are three ubiquitous initial/boundary conditions which arise in physics.

- The *Cauchy condition* specifies the values of ϕ and its derivatives. If ϕ is a function of n variables, the surface S in the parameter space should have dimension $n-1$, and if the PDE is order k, the Cauchy initial conditions must specify the values of ϕ and its first $k-1$ derivatives along the normal to the surface S. The system of the vibrating string above is an example of a Cauchy problem. We have a second order PDE with boundary conditions of sufficient order.

- The *Dirichlet condition* gives the value of ϕ on the boundary of the region of interest.

- the *Neumann condition* gives the value of the derivative at the boundary of the region of interest.

These boundary conditions are necessary to find solutions which are unique and well-behaved.

Let's look at an example of a Cauchy problem in the context of the wave equation in one dimension. Again, some general initial conditions are given for the interval [-∞,∞], $\phi(x,0) = f(x)$ and $\frac{\partial\phi(x,0)}{\partial t} = g(x)$. Using the method of characteristics (i.e. reduce a PDE to a family of ODEs), we can look for a solution as a sum of left and right going waves, expressed as:

$$u(x,t) = u_l(x+ct) + u_r(x-ct)$$

and so from our Cauchy conditions we have

$$f(x) = u_l(x) + u_r(x) \quad \text{and} \quad g(x) = c(u_l'(x) - u_r'(x))$$

After we differentiate the first equation and do some algebra we get

$$u_l'(x) = \frac{1}{2}(f'(x) + g(x)/c) \quad \text{and} \quad u_r'(x) = \frac{1}{2}(f'(x) - g(x)/c)$$

integrate these

$$u_l(x) = \frac{1}{2}f(x) + \frac{1}{2c}\int_0^x g(x')dx' + C_1 \quad \text{and} \quad u_r(x) = \frac{1}{2}f(x) - \frac{1}{2c}\int_0^x g(x')dx' + C_2$$

where $C_1 + C_2 = 0$ satisfies $u_l(x) + u_r(x) = f(x)$. Finally, our solution is simply

$$u(x,t) = \frac{1}{2}(f(x+ct) + f(x-ct)) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(x')dx'$$

Hence, all that's left to do is plug in the initial conditions. This particular case is called D'Alembert's solution.

7.5 Problems

1. Find the temperature distribution in a rectangular plate 10 cm by 30 cm if two adjacent sides are held at 100° and the other two sides at 0°.

2. Find the steady-state temperature distribution in a rectangular plate covering the area $0 < x < 10$, $0 < y < 20$, if the two adjacent sides along the axes are held at temperatures $T = x°$ and $T = y°$ and the other two sides at 0°.

3. A particle of mass m and momentum p is incident from the left on a potential step of value V_0. Calculate the probability that the particle is scattered backward by the potential if (a) $p^2/2m < V_0$ and (b) $p^2/2m > V_0$.

8 Complex Analysis

Complex analysis is very useful throughout physics. It is used in electrodynamics where the convention that the physical electric and magnetic fields are obtained by taking the real parts of complex quantities, quantum mechanics where solutions to the time dependent Schrodinger equation have an imaginary part, and in many-body physics where we use the residue theorem to find the Matsubara frequencies.

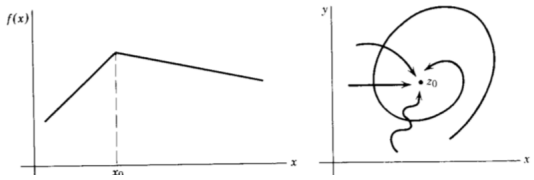
8.1 Cauchy-Riemman Conditions

$$f(z) = f(x + iy) = u(x, y) + iv(x, y)$$

Above is a complex function with real and imaginary parts. Usually functions are single-valued, that is $f(z)$ usually has one value for each z. Some complex functions are multivalued like $\ln(z)$.

$$\ln(z) = \ln|z| + i(\theta + 2n\pi)$$

where $\tan(\theta) = y/x$. For each z, $\ln(z)$ has an infinite set of values. However if θ is allowed a range of only 2π then ln (z) has one value for each z and this single-valued function is called a branch of z. The function $f(z)$ is analytic in a region of the complex plane if it has a unique derivative at every point of the region.



Above is an example of an analytic function and function which not analytic, there are an infinite number of ways to approach z_0 on the plot on the right and the derivative has a different value depending on the direction we approach. This leads to the Cauchy-Riemman conditions which are:

If $f(z) = u(x, y) + iv(x, y)$ is analytic in a region then in that region

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$$
$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$$

If $f(z)$ is analytic in a region then it has derivatives of all orders at points inside the region and can be expanded in a Taylor series about any point z_0 inside the region. The power series converges inside the circle about z_0 that extends to the nearest singular point.

8.2 Cauchy's integral theorem

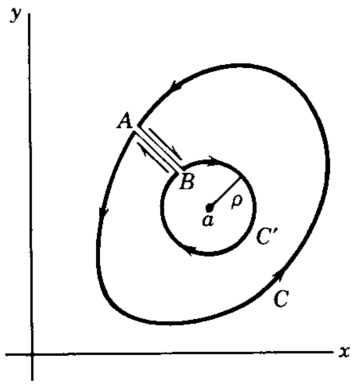
You can also do integrals on the complex plane. Cauchy's Theorem states the following:
Let C be a simple closed curve with a continuously turning tangent except possibly at a finite number of points. If $f(z)$ is analytic on an inside C, then

$$\oint_{\text{around } C} f(z)dz = 0$$

Cauchy's integral formula states: If $f(z)$ is analytic on and inside a simple closed curve C, the value of $f(z)$ at a point $z = a$ inside C is given by the following contour integral along C.

$$f(a) = \frac{1}{2\pi i}\oint \frac{f(z)}{z-a}dz$$

Cauchy's integral formula can be proven by using the following contour.



and taking $\rho \to 0$. Cauchy's integral formula is very useful and shows up in many areas in physics. For example it shows up in many-body physics when we want to evaluate Matsubara frequencies. For example the function $\frac{1}{e^{\beta z}+1}$ where $\beta = \frac{1}{kT}$ shows up in many-body physics and we want to evaluate it's simple poles. The function has simple poles when $e^{\beta z} + 1 = 0$. So $z = z_n = i(2n+1)\frac{\pi}{\beta} = i\omega_n$

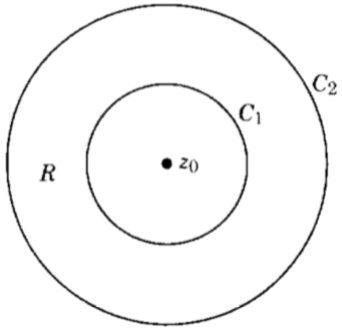
8.3 Laurent series

There is an analogue to Taylor series in complex analysis called the Laurent series which we can get using Laurent's theorem. Let C_1 and C_2 be two circles

with center z_0. Let $f(z)$ be analytic in the region R between the circles. Then $f(z)$ can be expanded in a series of the form

$$f(z) = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots + \frac{b_1}{z - z_0} + \frac{b_2}{(z - z_0)^2} + \cdots$$

convergent in R The "b" series is called the principle value of the Laurent series. We expect the Laurent series to converge inside the ring R since the "a" series is a power series and a power series converges inside C_2 while the "b" series is series of inverse powers of z, and so converges for $1/|z| <$ constant. Thus the "b" series converges outside some circle,C_1 and thus the Laurent series converges in R.

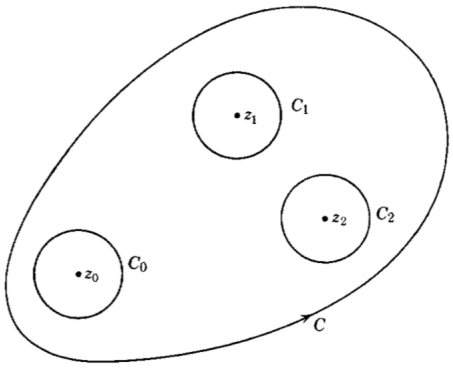


8.4 Residue Theorem

Let z_0 be an isolated point of $f(z)$. Now finding the value of $\oint_C f(z)dz$ around a simple closed curve C surrounding z_0 but inclosing no other singularities. Let $f(z)$ be expanded in a Laurent series about $z = z_0$ that converges near z_0 the only term that survives the integration is b_1 and is therefore called the residue.

when you have several isolated singularities we draw small circles as above and this leads to the residue theorem which states.

$$\oint_c f(z)dz = 2\pi i \cdot \text{sum of residues of } f(z) \text{ inside } C$$



8.5 Problems

1. Integrate $\int_{-\infty}^{\infty} \frac{s \cdot sin(sr)}{k^2 - s^2}$ using Cauchy's integral formula.
2. Boas 14.7.2

9 Special Functions

9.1 Legendre Polynomials

The Legendre differential equation is:

$$(1 - x^2)y'' - 2xy' + \ell(\ell + 1)y = 0$$

where ℓ is a constant. The most useful solutions of the differential equation are the Legendre polynomials. They first few Legendre polynomials are

$$P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{1}{2}(3x^2 - 1)$$

The generating function for the Legendre Polynomials is

$$\Phi(x, h) = (1 - 2xh + h^2)^{-1/2} \qquad |h| < 1$$

Taylor expanding the above function in term of $y = 2xh - h^2$ we get

$$\Phi(x, h) = P_0(x) + hP_1(x) + h^2P_2(x) + \cdots = \sum_{\ell=0}^{\infty} h^\ell P_\ell(x)$$

Some Legendre polynomial identities are

$$\begin{aligned}
&\ell P_\ell(x) = (2\ell-1)xP_{\ell-1}(x) - (\ell-1)P_{\ell-2}(x)\\
&xP'_\ell(x) - P'_{\ell-1}(x) = \ell P_\ell(x)\\
&P'_\ell(x) - xP'_{\ell-1}(x) = \ell P_{\ell-1}(x)\\
&(1-x^2)P'_\ell(x) = \ell P_{\ell-1}(x) - \ell x P_\ell(x)\\
&(2\ell+1)P_\ell(x) = P'_{\ell+1}(x) - P'_{\ell-1}(x)
\end{aligned}$$

The Legendre polynomials are orthogonal hence

$$\int_{-1}^{1} P_\ell(x)P_m(x)dx = 0 \qquad \text{unless } \ell = m$$

9.2 Problems

Boas 12.10.10

9.3 Bessel's equation

$$x^2y'' + xy' + (x^2 - p^2)y = 0$$

where p is a constant and called the order of the Bessel function.

$$\begin{aligned}
J_p(x) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+1)\Gamma(n+p+1)} \left(\frac{x}{2}\right)^{2n+p}\\
J_{-p}(x) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+1)\Gamma(n-p+1)} \left(\frac{x}{2}\right)^{2n-p}
\end{aligned}$$

are the two solutions of Bessel's equation. Any combination of these two solutions is also a solution of the Bessel function. One combination is called the Neumann function

$$N_p(x) = \frac{cos(\pi p)J_p(x) - J_{-p}(x)}{sin(\pi p)}$$

$$\frac{d}{dx}[x^p J_p(x)] = x^p J_{p-1}(x)$$
$$\frac{d}{dx}[x^{-p} J_p(x)] = -x^{-p} J_{p+1}(x)$$
$$J_{p-1}(x) + J_{p+1}(x) = \frac{2p}{x} J_p(x)$$
$$J_{p-1}(x) - J_{p+1}(x) = 2J'_p(x)$$
$$J_{p-1}(x) - J_{p+1}(x) = 2J'_p(x)$$
$$J'_p(x) = -\frac{p}{x} J_p(x) + J_{p-1}(x) = \frac{p}{x} J_p(x) - J_{p+1}(x)$$

$J_p(x)$ and $N_p(x)$ which are called Bessel functions of the first and second kinds. Since Bessel's equation is second order we have two independent but we there are also other functions which are also called Bessel functions. For example Hankel functions or Bessel functions of the third kind.

$$H_p^{(1)}(x) = J_p(x) + iN_p(x)$$
$$H_p^{(2)}(x) = J_p(x) - iN_p(x)$$

Spherical Bessel functions are closely related to Bessel functions of half odd integral order. That is if $p = n + 1/2$, n is an integer and $J_p(x), N_p(x)$. The spherical bessel functions are defined as:

$$j_n(x) = \sqrt{\frac{\pi}{2x}} J_{n+1/2}(x)$$
$$y_n(x) = \sqrt{\frac{\pi}{2x}} Y_{n+1/2}(x)$$
$$h_n^{(1)}(x) = j_n(x) + iy_n(x)$$
$$h_n^{(2)}(x) = j_n(x) - iy_n(x)$$

9.4 Spherical Harmonics

If the potential in the three dimensional Schrodinger equation only depends on the distance from the origin $V(r)$ then it is natural to use spherical coordinates (r, θ, ϕ). We can then separate it into a radial equation and a angular equation.

The angular equation solution are the spherical harmonics:

$$Y_\ell^m(\theta,\phi) = \epsilon\sqrt{\frac{(2\ell+1)}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}}e^{im\phi}P_\ell^m(cos\theta)$$

where $\epsilon = -(1)^m$ for $m \geq 0$ and $\epsilon = (1)$ for $m \leq 0$. P_ℓ^m is the associated legendre function. Defined as:

$$P_\ell^m(x) = (1-x^2)^{|m|/2}\left(\frac{d}{dx}\right)^{|m|}P_\ell(x)$$

and $P_\ell(x)$ is the ℓth Legendre polynomial defined as:

$$P_\ell(x) \equiv \frac{1}{2^\ell \ell!}\left(\frac{d}{dx}\right)^\ell (x^2-1)^\ell$$

9.5 Problems

Show that $\int_0^\infty J_1(x) = -J_0(x)|_0^\infty = 1$

Show that

$$\int_0^{2\pi}\int_0^\pi Y_\ell^m(\theta,\phi)^* Y_{\ell'}^{m'}(\theta,\phi)sin(\theta)d\theta d\phi = \delta_{\ell\ell'}\delta mm'$$

References

[1] *Boas, M. L. (1983). Mathematical methods in the physical sciences. New York: Wiley.*

[2] *Taylor, John R. (John Robert), 1939-. (2005). Classical mechanics. Sausalito, California: University Science Books*

[3] *Griffiths, D. J. (1999). Introduction to electrodynamics. Upper Saddle River, N.J: Prentice Hall.*

[4] *Griffiths, D. J. (1995). Introduction to quantum mechanics. Upper Saddle River, N.J: Prentice Hall.* Nearing *James Nearing (2004). Mathematical Tools for Physics.*

[5] *Lim, Yung-Kuo (1991). Problems and Solutions on Optics. World Scientific*

[6] *Lim, Yung-Kuo (1998). Problems and Solutions on Quantum Mechanics. World Scientific*