Physics Notes: Calculus Based (8/8/2026 DRAFT COPY)

Contents

Part I

Introduction and Mathematics

Chapter 1

Introduction

1.1 Statement of Purpose and Credits

These notes can be thought of as a continuation of the ones from AP physics 1. For brevity, I won't repeat things that were in those notes. You can go read them if you need to review. If anything doesn't flow, or is confusing, please let me know so I can (eventually) improve.

1.2 Useful Outside Materials

- Textbooks:
  - Griffith: Introduction To Electrodynamics

    Somewhat more advanced than this course, but much of it is relevant. Probably the most common text in university E+M courses.
  - Purcell: Electricity and Magnetism

    My personal favorite undergrad textbook. Somewhat more advanced than this course. Very enlightening explanations in terms of thought experiments. Don't worry about anything related to relativity...
  - Halliday and Resnick: Fundamentals of Physics

    An intro textbook that covers all the material in the class, and a lot more. I have some issues with their high degree of reliance on formulas and lack of derivations, but if you want a textbook at the AP C level, this is probably the best choice.
- Simulations: I strongly recommend looking into a simulation if you are struggling to grasp something. The ability to visualize can really help.
  - ophysics.com: excellent resource for much of what is covered here.
  - phet.colorado.edu: has a few good resources. The level of the material varies from middle school to second year of college, and finding relevant stuff can be a chore

1.3 Image Credits

All of the figures are my own creations unless noted otherwise.

1. Figures were made with some combination of Mathematica, Geogebra, EDrawMax and Tikz (hence the very different looking images).

Chapter 2

Mathematical Prerequisites

In this chapter some of the mathematical structures that we will use this year will be laid out. You don't necessarily need to be an expert in them immediately since we will refer back to them a number of times over the course of the year, but you should be fluent enough with them to apply them to simple problems. If you have forgotten any of the basics of what vectors are or how to use them, refer back to the physics 1 notes. In particular, if you don't know the dot and cross product by heart, you are going to immediately run into problems in this class.

I intentionally am not providing detailed derivations of some things. You can find correct derivations of everything in a calculus based physics class in any good undergraduate textbook. For some reason it is very hard to find anyone explaining intuitively what the math means or why we are doing it. That is the purpose of these notes.

2.1 Scalar Fields

When someone tells you that it is $40°C$ outside, you have some understanding of what that means, but with some thought, it should become far less obvious what the meaning of that statement is. There isn't one temperature outside. Instead, temperature is a function of your position on the ground as well as your altitude and time so we really should be specifying this function $T(x, y, z, t)$, which we call a scalar field. This just means that we have a scalar defined at every point in space and time. We won't usually be concerned with time variance, and we will simplify the notation using vectors, so our scalar temperature field could be written as $T(\vec{r})$.

We could in principle place thermometers at every possible location, and each would read the current local temperature. This is obviously quite impractical, so what we would do in practice is measure the temperature at convenient locations and use the property of continuity, that is that two values of the scalar field taken from nearby points should be similar in value. More concretely, it is highly unlikely to be 0°C in your front yard, while also being 40°C in your back yard.

Ultimately, the concept of a scalar field doesn't provide us with much extra utility in this simple example because we have no way to measure or calculate the value, but now lets consider a different scalar field: gravitational potential energy per unit mass above the surface of Earth. We will simplify the discussion by considering only relatively small heights so that we can use that $U(h) = mgh$ gives the potential energy of a mass at height h.

The potential energy per unit mass is $V(h) = gh$. In this case, we did gain something from the idea of a field, we can now simply multiply the scalar field $V(h)$, by whatever mass we are interested in to get the potential energy of that mass at any arbitrary height. It may not seem like we got anything new from this, since we always calculated the potential energy of a mass using mgh. The new insight is that we were only

able to do this because gravity is a scalar field dependent only on height. If no such field existed, calculating gravitational energy would be much harder.*

2.2 vector fields

A natural question to ask after defining a scalar field is whether this could be extended to include vector quantities. The answer is yes. In our gravity example, the gravitational acceleration $\vec{g}$ is a scalar field with a constant magnitude and a direction that always points "down". In this case it is worth considering what down even means. Shouldn't our "down" be Australia's "up"? The answer is yes... if we insist on using Cartesian coordinates. Let's consider the problem instead using spherical coordinates. The $\hat{r}$ unit vector in spherical points from the origin to a point. For our case we will use the center of Earth as our origin. This means that $\hat{r}$ for any observer now always points straight from the center of earth to that observer. Now Australia and the USA agree. The gravity field is $\vec{g} = -g\hat{r}$ in other words a magnitude of g with a direction that points towards the center of Earth.

Strictly speaking, we can only say that we have a vector or scalar field when a vector or scalar is defined at every point in a domain. Practically, vector fields will often be defined even when we haven't actually made an appropriate definition at every point. If you were planning shipping routes, you could imagine being interested in the velocity of water in the river. The resulting map would only have a finite number of points (because measuring everywhere would take literally forever). The resulting object that comes from plotting all these vectors together would still usually be called a vector field.

2.3 Cylindrical and spherical coordinates

In previous classes we have expressed positions by giving an x, y, and z coordinate. This results in position vector of the form $\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}$. Unfortunately, it will turn out that most of the objects that physics is interested in studying are cylinders/disks (galaxies, the solar system, car wheels...) or spheres (planets, stars, golf balls...). It is not impossible to describe spheres or cylinders in Cartesian coordinates. but it also isn't very convenient.

Fortunately, you already know the solution. On earth, we can give latitude, longitude and elevation rather than x, y, z. These are three independent coordinates, so they uniquely specify a location. Note that these are two angles and 1 distance. In the case of height, distance is set to be 0 on the surface of Earth. This isn't very convenient for anything except Earth, so we will set 0 to be at the origin (often the center of the sphere). We can thus replace Cartesian (x, y, z) with spherical r, θ, ϕ. Here ϕ is like longitude (ie it measures position along the equator, or East-West position) and θ is the **opposite** of latitude, ie it is 0 at the poles and $\frac{\pi}{2}$ at the equator. Equivalently, and more usefully in practice, $\hat{\theta}$ is the angle from the positive z axis, and $\hat{\phi}$ is the angle from the positive x axis. r represents the distance from the origin to that point.

The use of spherical coordinates can make describing some objects much simpler. As a basic example, consider a spherical shell with a radius R. In Cartesian coordinates, we could say that the equation of this sphere is $x^2 + y^2 + z^2 = R^2$. But notice that $x^2 + y^2 + z^2$ is just the square of the vector that points from the origin to a point on the shell. Spherical coordinates notes that $\sqrt{x^2 + y^2 + z^2} = R$ is just a way of writing that every point is the same distance, R, from the center of the sphere. So it would be very nice if we just took the coordinate r, then the equation of the spherical shell is just $r = R$, which is much easier to work with.

But what about directions? We let $\hat{r}$ be the unit vector that points from the origin towards the object. That definition feels pretty natural, but what about the two angles? We define $\hat{\theta}$ to be perpendicular to $\hat{r}$ in the plane defined by θ. This is really abstract, look at figure 2.1 to help with understanding here. Note that the three unit vectors are always perpendicular to each other, but they point in different directions

*We will see just how much harder when we talk about magnetism, for which no scalar field may be defined.

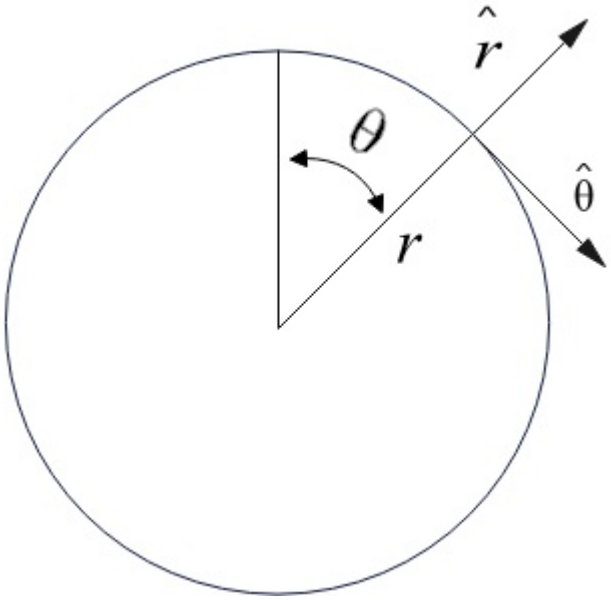


Figure 2.1: $\vec{r}$ will point from the origin to the point on the circle where the unit vectors $\hat{\text{r}}$ and $\hat{\theta}$ are shown. This is a distance r from the center of the circle. Note that $\hat{\text{r}}$ and $\hat{\theta}$ are perpendicular. In this diagram $\hat{\phi}$ would point straight into the page.

depending on where we are on the circle. This, again, is familiar. Depending on where we are on Earth, saying "walk East" could represent walking any direction in a Cartesian system. If you start in Asia, you could even sail East to end up West of where you started. This is the main downfall of spherical coordinates, and will make it nearly impossible to integrate vectors in spherical coordinates (because the unit vectors are time dependent). Integrating scalars though, is often much easier.

Cylindrical coordinates keeps Cartesian z, but replaces x and y with $r = \sqrt{x^2 + y^2}$ and ϕ.[†]

Consider an object moving around a circle with radius R with some angular speed ω. We could write its position in Cartesian coordinates as $\vec{r} = R\cos\omega t\ \hat{\text{x}} + R\sin\omega t\ \hat{\text{y}}$ (remember that circular motion is just sine motion in one direction and cosine in the other). By differentiation, we could say that $\vec{v} = -\omega R\sin\omega t\ \hat{\text{x}} + \omega R\cos\omega t\ \hat{\text{y}}$.

In cylindrical, to get position we would just note that it is always wherever it was, since we defined $\hat{\text{r}}$ to point from the origin to its location we would just say that $\vec{r} = R\hat{\text{r}}$. For the velocity, we would just note that they are always moving tangent to the circle (or equivalently perpendicular $\hat{\text{r}}$) which is how we defined $\hat{\phi}$. This means that its velocity should be $\vec{v} = \omega R\hat{\phi}$.

2.4 Integrals in Multiple Dimensions

2.4.1 What is an Integral?

You probably know that an integral gives the area under the curve of some function. This is the best kind of statement in that it is technically correct, but functionally useless. Lets think about how we can build some intuition for what an integral is really doing. The essence of the idea is that an integral takes an infinite set of objects of dimension n and returns an object of dimension $n + 1$. In general, the dimensions might be

[†]Notation is nowhere near consistent here. In practice you will see θ used here just as often as ϕ. I prefer ϕ because then ϕ represents an angle in the $x - y$ plane in both cylindrical and spherical.

something very weird, but here we talk only about area and volumes.

Let's say we want to create a square with side length L out of lines with length L oriented along the y axis. If we wanted to, we could (in principle) stack an infinite number of these infinitely thin lines on top of each other to get something with a finite height. Assuming that the line has thickness of dx, we could perform the integral

$$A = \int_0^L L \; dx = L^2$$

which is the area of our desired square. If we wanted to turn this into a cube, we could use that we now have a square with sides of length L and thickness dz. So we integrate this along z to get

$$V = \int_0^L L^2 \; dz = L^3$$

which is the volume of our desired cube.

You might object to this by claiming I chose cubes because the constants are all 1. So let's do a disk. We'll even start with points instead of rings. First, we want to make a ring of radius r. To do that we take a point with thickness dr some distance r from the origin and revolve it through an angle of 2π. Then our integral is

$$A_r = \int_0^{2\pi} r \; dr \; d\phi = 2\pi r \; dr$$

Which is the area of a thin ring with radius r and thickness dr. If we want to add up a bunch of rings with different r, (but the same thickness dr), we integrate

$$A = \int_0^r 2\pi r \; dr = \pi r^2$$

Which is of course the area of our desired disk. You can check for yourself that this works for spheres, or any other shape.

Now imagine we need to weight our points, possibly because a point that has a higher density should count more in determining the mass. For finite sums, that is pretty easy to do. We could just take the sum

$$\sum_i W_i x_i$$

where W_i is how much weight we wanted to give x_i. Applications for this might be something like finding the average score for a class. In that case x_i would be a score, and $W_i$would be the number of students who received that score. To get the average we would need to divide by the number of students, so we would get

$$< x >= \frac{\sum_i W_i x_i}{\sum_i W_i}$$

We can weight integrals in the same way, but now our weights are replaced by some weighting function f

$$< x >= \frac{\int x f(x)}{\int f(x)}$$

If you really understood this, then the extension to multi dimensional integrals is actually pretty straightforward.

2.4.2 Volume and Area Integrals

To find the volume of our cube, instead of starting with a line of length L and integrating twice, we could have just started with points. I like visualize integration by thinking about what it would look like for us to trace out all the points in our object. For our cube, we need to trace out every point in a line along x, then drag that line through y and finally drag the square through z (in Cartesian only, you could also think in terms of multiplication of dimensions. Don't.). This tells us the volume element in Cartesian is just $dV = dx\ dy\ dz$

$$V = \int_0^L \int_0^L \int_0^L dx\ dy\ dz$$

Now lets think about cylindrical. For cylindrical, z is the same as Cartesian. The new interesting bits are r and ϕ. If I vary ϕ at a given r, the amount of new material I would add to my object depends on how big r is because I would be tracing the circumference of a circle with radius r. This means that I need to add an extra term to my integral to account for the dependence on r. Equivalently, the volume element in cylindrical coordinates is

$$\boxed{dV = r\ dr\ d\phi\ dz} \tag{2.1}$$

$$V = \int_0^h \int_0^{2\pi} \int_0^R r\ dr\ d\phi\ dz$$

Typically each variable is a constant with respect to other variables, so it can be safely separated. Note that this was more general than it appeared. If we want to integrate things in cylindrical, we just have to remember to add that extra factor of r inside the integral. For example if we want the mass of a cylinder with some $\rho(r, \phi, z)$ the integral is just

$$M = \int_0^h \int_0^{2\pi} \int_0^R \rho(r, \phi, z)\ r\ dr\ d\phi\ dz$$

Spherical is substantially trickier because we have 2 angles, and they are functionally different. As seen in figure 2.1, θ is the angle from the North pole, and ϕ is the angle around the equator. If you have a globe, referring to it might help here. Otherwise, you will need to imagine the earth.
Take some small fixed angle θ. Now rotate all the way around in ϕ. How far did a point travel as you did that? Not very far... Generally, you will quickly see by playing with it, that the larger that θ is, the larger distance you move as you vary ϕ by the same amount. If you look at 2.1 again, you should be able to see that the size of a circle carved out as ϕ is varied is proportional to $\sin\theta$.

More precisely, if we increment by $d\phi$, then we will go a distance of $r\sin\theta d\phi$. Now we need to increment θ. When we increment θ by some tiny amount, the distance we move is just $r\ d\theta$, by the same argument we used for ϕ in polar coordinates. Putting that together, that tells us that an infinitesimal volume element in spherical coordinates is

$$\boxed{dV = r^2 \sin\theta\ dr\ d\theta d\phi} \tag{2.2}$$

$$V = \int_0^{2\pi} \int_0^{\pi} \int_0^R r^2 \sin\theta\ dr\ d\theta\ d\phi \tag{2.3}$$

Analogously to cylindrical, we basically just need to modify our integrals by adding $r^2\sin\theta$ to the integrand.
If all this felt hard, try finding the mass of a sphere with radius R and density given by $\rho = Ar$ in Cartesian... You'll be there a while...

2.4.3 Surface Integrals

In 1D, the definition of integral was fairly unambiguous (even if calculating them was usually impossible). In multiple dimensions, different types of integrals exist. For example, imagine I want to know how much water is flowing through a hula hoop that I place in a river. I should be able to find this by breaking up the region of interest into very small regions and adding up ρv from all of those small regions over the whole area (note that $\rho A v$ has units of $\frac{kg}{s}$ so this is the rate that mass is being carried through). Since ρ is constant, it seems like I just need to measure v everywhere and add them all up. This sounds pretty doable, but unfortunately is wrong.

To see why, imagine rotating the hoop so that it is parallel to the flow of the water rather than perpendicular. The integral now is very easy. No water flows through. This means that what we cared about wasn't just speed. Instead we had to care about the direction of the velocity and how it relates to the orientation of the hoop.

Fortunately, we already have an operation that asks how much of one vector is parallel to another: the dot product. Clearly one vector in the product should be the velocity, but the other isn't as clear. To get around this, note that the amount of material passing through the loop is maximized when the plane of the loop is perpendicular to the velocity and decreases as it gets closer to parallel as we can see in figure 2.2.

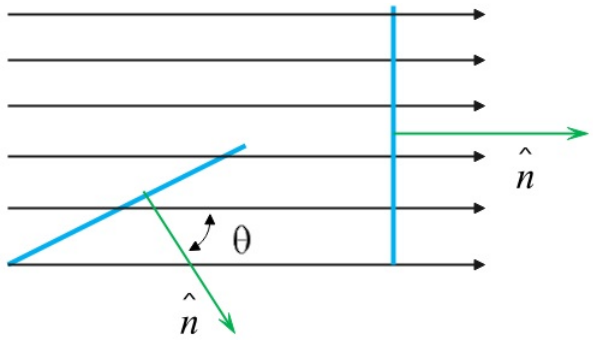


Figure 2.2: Two surfaces with the same area placed in a river (yes, the blue lines are the same length, optical illusions are fun). Notice that the flux through the surface oriented perpendicular to the flow has a larger flux through it.

We typically call the unit vector perpendicular to a surface $\hat{n}$. In fig 2.2, the green arrows are $\hat{n}$.[‡] So if our velocity is constant, then the rate at which mass crosses the surface is

$$\Phi_m = \rho \vec{v} \cdot \hat{n} A$$

Note that if $\vec{v}$ is parallel to $\hat{n}$ then this just simplifies to scalar multiplication and gives the answer from the previous paragraph. If $\vec{v}$ or ρ aren't constant across the hoop then we can't simply multiply. Instead, we'll have to perform the integral:

$$\Phi_m = \iint \rho \vec{v} \cdot \hat{n} \, dA$$

The two integral signs mean that this is an area integral , so we are integrating over 2 separate variables (in Cartesian, this might be x and y). Whether to use one or two integral signs here is a matter of notation. Most sources prefer one with the logic that A is the variable of integration, not whatever coordinates give A. Also, remember that $\vec{v} = \vec{v}(\vec{r})$: it depends on location in space. So does $\hat{n}$ generally, since the surface is

[‡]An astute reader may notice that there are two vectors that could work for $\hat{n}$. They point in opposite directions. In our example no convention really exists. I just take the one that produces positive dot product unless there is a good reason not to. In 3D, the convention is to take the outward facing normal vector.

probably curved. As you might imagine, this integral is very hard to evaluate in general, the only time it will make sense to evaluate surface integrals is when there is a very high degree of symmetry to simplify the calculation. This integral is usually written less explicitly as

$$\Phi_m = \int_S \rho \vec{v} \cdot d\vec{A} \tag{2.4}$$

where we just use that $d\vec{A}$ is a shorthand for $\hat{n}dA$ and noted that the fact that we are integrating for area should be clear by the dA, so we don't need the second integral sign. The S in the integral bounds just reminds us that the area elements dA that we are integrating over are taken from the surface S.

I don't like removing $\hat{n}$, but that is the default on the AP. I do like removing the second integral sign, and specifying the surface S. The AP uses a single integral, but does not specify the surface explicitly. Either way, you will see many notations here... get used to it.

This example was done with fluids, but the ideas are general. If we are concerned with the flux of a vector field $\vec{S}$ through a surface, we would just have

$$\boxed{\Phi_s = \int_S \vec{s} \cdot \hat{n} \; dA} \tag{2.5}$$

We will discuss surface integrals many times in this class, some examples are gravity and the electric field.

2.4.4 Line Integrals

A line integral integrates the dot product of two vectors along some curve. This sounds weirdly abstract, so let's use the familiar concept of work.

Energy can be transferred into or out of the system by doing work on it. We will take the definition of work to be

$$W = \Delta E$$

we can hopefully recall that there were different ways to find work in a specific situation. For the case of friction, which always opposes motion, $W = FD$. For gravity, which always points in the same direction (for displacements much smaller than Earth's radius), $W = \vec{F}_g \cdot \Delta\vec{r}$. Imagine now that we had a force whose angle with our motion varied with time. Because the functions must be continuous, we can break the path into very short intervals where the angle between the force and the small displacement is approximately constant. Then we could sum up all those small intervals.

$$W = \sum_i \vec{F}_i \cdot \Delta\vec{r}_i$$

If we make these intervals infinitesimal, than we get an integral:

$$W = \int_C \vec{F} \cdot \vec{dr}$$

Where the C in the integral is not a bound, instead it is a statement that the integral is to be taken along some curve (the path taken by the particle in this case.)

Once again, this integral is only possible if the problem has very nice symmetry. The only new case we will consider in this class is the case where we travel along some straight line with a force that is explicitly a function of position or time. The most interesting case is where we have some time dependent force acting on an object initially at rest.

The integral above looks pretty intimidating, but with our simplifications we can write

$$W = \int_{t_1}^{t_2} \vec{F} \cdot \frac{d\vec{r}}{dt}\, dt \tag{2.6}$$

or

$$W = \int_{t_1}^{t_2} \vec{F} \cdot \vec{v}\, dt$$

Equation (2.6) is actually pretty general. Often, the only way to evaluate a line integral is to find a way to parameterize the movement of the particle so that we can instead perform a simple(r) integral for time. We will discuss line integrals in details when we talk about energy and again when we talk about magnetic fields.

2.5 Derivatives in Multiple Dimensions

Derivatives in 1-D were easy to define because only one rate of change made sense. Contextually that might be a derivative with respect to t, or with respect to x, but it was clear from context. In multiple dimensions, you must be told what other variables are doing before calculating the rate of change with respect to some variable. As a quick example, say we want to find the rate of change of the temperature. We might write $\frac{dT}{dt}$. The problem happens if you are comparing between observers who are moving and stationary. We expect that if you are decreasing in altitude, you would measure a much more rapid temperature change than someone at rest. We thus need to define some new derivative that explicitly keeps every other variable constant. We denote the derivative of some function $T = T(x, y, z, t)$ with respect to x, keeping y and t constant

$$\frac{\partial f}{\partial x}$$

and we reserve the use of the traditional d for derivatives for the case where we want a total derivative. Normally it only makes sense to take total derivatives with respect to time.

Let's simplify by taking out time (imagine we are looking at data of temperatures across Phoenix all taken at the same time). We don't care about the curvature of Earth since Phoenix is really small by comparison. Thus we can use Cartesian and the temperature function becomes $T = T(x, y, z)$. We want to know what direction we can go to find the hottest nearby point (finding the hottest point globally is a very different problem).

In 1D, if we wanted the value of a function f at a **nearby** we could imagine moving some small distance Δx. Then we use some logic to see that the new value should be the value at the starting point $f(x_0)$ added to the change Δf. To get the change we multiply the rate of change of the function with respect to x (ie $\frac{df}{dx}$), by the distance we moved Δx. Putting this together in math

$$f(x_0 + \Delta x) = f(x_0) + \frac{df}{dx}\Delta x$$

Note that this is equivalent to saying we take a step along the tangent line. Obviously this only works if our displacement is small enough that the tangent line doesn't differ meaningfully from the function. If we make Δx small enough, this should work for any function that doesn't have discontinuities[§].

Intuitively, we should move along the direction of fastest increase, which will in general be some line that doesn't point along any of our axes. If the temperature increases rapidly in $+\hat{\text{x}}$, increases slowly in $-\hat{\text{y}}$, but decreases both along $+\hat{\text{z}}$ and $-\hat{\text{z}}$, it makes sense that we should move along a line that is between $+\hat{\text{x}}$ and $-\hat{\text{y}}$ with no component along $\hat{\text{z}}$. The rate of change in the $\hat{\text{x}}$ direction is $\frac{\partial T}{\partial x}$, with similar formulas for $\hat{\text{y}}$ and $\hat{\text{z}}$.

[§]With the exception of $\frac{1}{r^2}$ forces at the origin, discontinuities in physics equations are extremely rare

Now we want to generalize the math to three dimensions. Superposition does the trick for us, we use that the total change from movement is the sum of the individual changes from movement in each direction (note: this only works if our directions are perpendicular, but you should chose ones that are...).

For our function T, this becomes

$$T(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) = T(x_0, y_0, z_0) + \frac{\partial T}{\partial x}\Delta x + \frac{\partial T}{\partial y}\Delta y + \frac{\partial T}{\partial z}\Delta z$$

Naively we might guess that we should just move along the direction with the largest partial derivative. This is wrong. See 2.7 for a proof of this).

Our problem is that the actual fastest increase direction is probably some line that doesn't point along any of our coordinates. We thus seek an operation that gives us a vector pointing in the direction of fastest increase. Taking our cue from the 1D case, we should probably try something like the tangent vector, but that is a dead end. To get to our goal, we define a new thing called "del" and represented with the symbol $\vec{\nabla}$. It is a vector operator defined (in Cartesian coordinates) as

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

Note that this is an operator, something that acts on a function and returns a function, so we should usually not be writing it by itself. If the idea of an operator defined without a function seems weird, it really shouldn't. You didn't define the ordinary derivative in terms of any specific function. The ordinary derivative behaves the same way on any function, so the choice to put a function after it when defining it was just a crutch to make it more familiar to you. The added generality of an abstract definition wasn't needed for 1D, but we will need it here because del can be applied to different objects and in different ways. The first meaning of $\vec{\nabla}$ that we will talk about is del applied to a scalar function. Let's use our temperature function.

$$\vec{\nabla}T = \frac{\partial T}{\partial x}\hat{\mathrm{x}} + \frac{\partial T}{\partial y}\hat{\mathrm{y}} + \frac{\partial T}{\partial z}\hat{\mathrm{z}}$$

This instantly produces a vector, but more importantly, it is a vector whose components are proportional to the rate of change in that direction. This means it points in the direction of fastest increase of the function. Be careful with this interpretation though. For a function $f(x, y)$, the vector is in the x,y plane. For our temperature function, it is a true 3D vector. The introduction of del solved our problem immediately. The result of applying del this way is typically called the gradient (or more properly gradient field) of the scalar function. See figure 2.3 to visualize this for some a function $f(x, y)$

Now our question of how to find the temperature at some nearby point reduces to

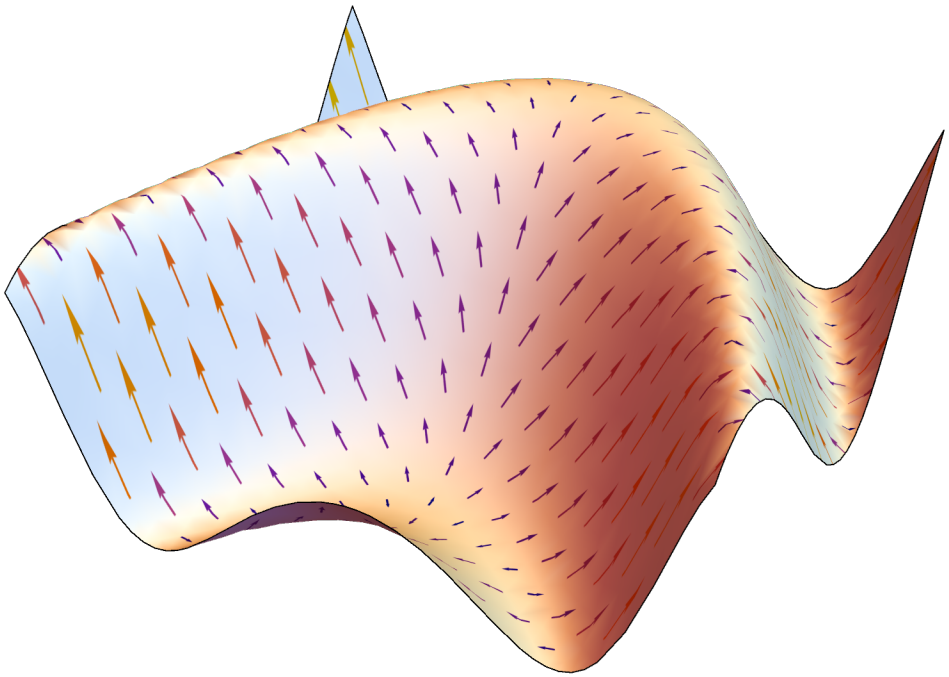
$$T(\vec{r} + \Delta\vec{r}) = T(\vec{r}) + \Delta\vec{r} \cdot \vec{\nabla}T$$

Which is way nicer than before. To get some idea of what is going on here, imagine that you moved a unit length in a direction perpendicular to the gradient. The temperature would not change. If you move along the gradient it changes as quickly as possible. If you move at some angle to the gradient, we would need to know how close to parallel to the gradient your motion was. Finding out how close two vectors are to parallel was the function of the dot product. Note also that if the gradient is longer, or your step is bigger, we would get a bigger increase as expected.

2.5.1 Divergence and Curl

Consider the two fields shown in figure 2.4. They see pretty obviously qualitatively different. One appears to only spread out from a point, while the other only circulates around it. You might imagine the pure

Figure 2.3: A plot of the function $f(x, y) = \sin(x + y^2)$ overlaid with the gradient field. The surface represents the value of f at those coordinates. Notice that the gradient always points in the direction where the function increases the fastest and that the length is proportional to how fast it increases there. A common point of confusion is thinking that the gradient vector lives in the 3D space. The gradient of a 2D function is 2D. It lives in the plane. Visualizing it with images like this make the relationship with the curve clearer at the expense of obfuscating the relationship with the plane.



spreading as being light from a star, and the pure circulation as being a whirlpool in water. We are looking for some way to quantify the idea of "spreads out" and the idea of "circulates around a point". We already have the tools to do this in the form of surface integrals and line integrals.

Remember that flux was a surface integral and can be thought of as how much a vector field penetrates a surface. Mathematically, for a closed surface

$$\Phi_s = \oint \vec{F} \cdot \hat{n} \, dA$$

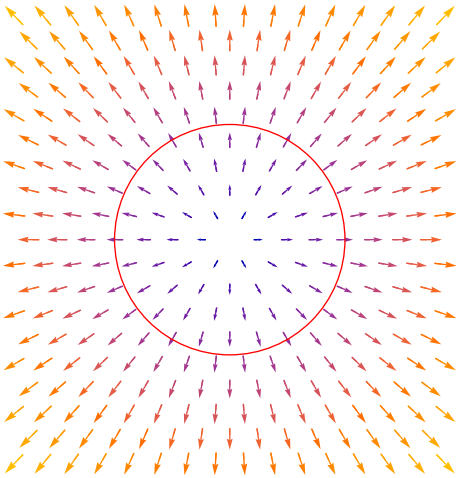
Each plot has a circle drawn around the the center of the field, you should immediately notice that the dot product of the field with the surface normal $\hat{n}$ is positive everywhere for 2.4a, so there is a positive flux through the surface. The field in 2.4b instead has the field always parallel to the surface, so the dot product is 0 everywhere and the total flux through the surface is 0.

By contrast,if we take a line integral of the vector field along the closed circle

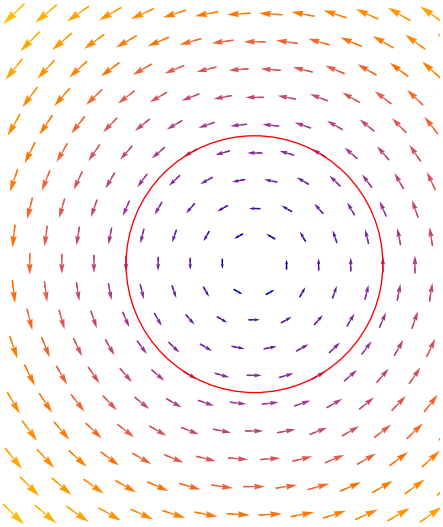
$$\oint_C \vec{F} \cdot d\vec{l}$$

the pure divergence field gives 0 because the field is never parallel to the tangent to the circle. The pure curl

Figure 2.4: Fields with divergence behave very differently than fields with curl.



(a) A field with large divergence but 0 curl. Important note: if the field was $\vec{f} = \frac{\hat{r}}{r}$, the field would look quite a lot like this, but the arrows would get smaller moving outwards. It turns out that field has nonzero divergence only at the origin.



(b) A field with large curl but 0 divergence. Note: the field $(-y/r^2, x/r^2)$ would have a similar shape, but non-zero curl only at the origin.

field gives something not 0 because the field is always parallel to the surface.

We can take our intuitive ideas to be

- divergence: a measure of how much outward flux is being sourced at a point
- curl: a measure of how much circulation is being sourced at a point

Note that both of these are defined at each point, not for the whole field at once. This is what gives rise to the weird results mentioned in 2.4. While the other fields mentioned certainly seem to circulate macroscopically, they don't exhibit microscopic circulation at a point. The fact that the field given by $\vec{f} = (\frac{\hat{r}}{r})$ fails to diverge in 2D is that the length of a circular arc gets longer by a factor of exactly r as we move away from the origin. This means that total amount of field in any circle around the origin is fixed for that field. Equivalently, nothing is being sourced or sinced, it is just spreading out in space, hence no divergence. In 3D, a $\frac{1}{r^2}$ field has the same property. The argument for curl is similar if a bit trickier. If this is all confusing, $\frac{1}{r^2}$ fields will be discussed in great detail when we talk about gravity and electrostatics.

Now we need to turn these concepts into math.

We'll start with divergence. What makes a field diverge? Intuitively the answer is that there is some source of the field. If we drew a surface around you and measured the CO_2 flux through the surface, we would get something nonzero because you are producing CO_2, so more would leave the circle than enter it. You are a source for carbon dioxide. This analogy is accurate, but hard to think about, so I am going to switch to an easier to consider analogy.

Consider a light-bulb If we drew a circle around the bulb, we would find that more photons left the circle than went in. That is, there was a net flux through the surface. The field must thus have nonzero divergence and be sourced by the light-bulb. Note that in this case, the divergence is 0 everywhere except the light-bulb

because light is not being created or destroyed anywhere else.

Now imagine that we have some fluid. If we are heating it in a microwave, it will be expanding slightly. This must mean that if we consider an (imaginary) sphere inside the fluid that was being heated, we would find that there was a net outward flux through the sphere. Note that an imaginary sphere in a uniform river would not produce this effect because all fluid that enters also leaves. We conclude that only creating more of something (in this case volume of fluid) contributes to divergence. If the river turns, the direction of the fluid velocity might change, but the total volume entering and leaving would be the same. Note that I am NOT saying that the total amount of fluid increases in our microwave example, decreasing the density results in increasing the volume.

Mathematically, in the straight river case, the derivatives in any direction were just 0. The turning river would have a negative derivative in one direction, but an equally large positive derivative in another. What we seek is thus an object that sums the derivative across all directions and returns some scalar function that tells us how much the field tends to diverge there. Let's try

$$\vec{\nabla} \cdot \vec{f} = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}$$

This does exactly what we want. We will call this the divergence.

Now we consider Curl. An astute reader might note that we already have scalar multiplication (gradient) and dot product(divergence), so there is one other operation that makes sense, the cross product. But does it do what we want? Recall the cross product gives us how much of a vector is perpendicular to something. If the dot product gave tendency to be perpendicular to a surface, the cross product could only give tendency to be parallel to it. So we have our third use of $\vec{\nabla}$

$$\vec{\nabla} \times \vec{f}$$

Which we call the curl of the field. I didn't write out the explicit definition in terms of coordinates because it is long and we won't use it a single time in this class, but we will use the concept of curl when we talk about magnetic fields. Also, it should be clear either from the use of the cross product, or from the idea that it needs to produce something that circulates but does not penetrate any surface that the curl should be taking derivatives of the component in each direction with respect to every OTHER direction. This is the opposite of what the divergence did.

2.5.2 The Divergence Theorem and Stokes Theorem

If we want to know how much flux a vector has through a closed surface, we used the surface integral. That must be related to divergence, which told us how much outward flux was being sourced at a point. Specifically if we sum the amount of flux being produced at any point over all points contained in a closed surface, that better give the flux through the surface. It does. This is one of two fundamental theorem of calculus equivalents in multiple dimensions. It is often called the divergence theorem.

$$\boxed{\oint_{\partial V} \vec{F} \cdot d\vec{A} = \int_V \vec{\nabla} \cdot \vec{F} \; dV} \tag{2.7}$$

Similarly, if we want to know how much total circulation there is along a circle like the one from fig 2.4b, we better be able to sum all the circulation sources (ie the curl) in the circles area to get the circulation along the perimeter. We can. This is the other fundamental theorem of calculus in multiple dimensions. It is called the Stokes Theorem.

$$\boxed{\oint_{\partial A} \vec{F} \cdot d\vec{l} = \int_A \vec{\nabla} \times \vec{F} \cdot d\vec{A}} \tag{2.8}$$

Congratulations. You now know all of multi-variable calculus.[¶]

[¶]This of course assumes you understood everything perfectly, and also only accounts for a typical first year multi-variable course.

2.6 Differential Equations

Several times in this class we will need to solve equations of the form

$$\frac{d}{dt}Y = f(Y,t)$$

A note on notation here. For historical reasons, physicists use a dot over a variable to denote time differentiation. This was Newton's notation. Every other field uses Leibniz's notation. I will use both.
In general, these equations do not have closed form solutions in terms of elementary functions (eg, trig functions, polynomials and exponentials), but a few special cases are easily solved. If you take a differential equations class, you will learn a huge kit of ways to solve various slightly different looking equations. I don't have time for that, so we will discuss only one techniques here. ‖

2.6.1 Ersatz Solution

The first method is called either un-determined coefficients, or an Ersatz solution, because math people don't like to call it guessing. The technique is that you make a clever guess involving a constant, then solve for the constant.
The most useful cases for this are of the form

$$\frac{d^n t}{dt^n} = Ay$$

where A is any combination of constants and n is any integer. For these we note that the only function that returns itself with a constant as a derivative is an exponential, so we guess

$$y = y_0 e^{\omega t}$$

Plugging this back into the equation gives

$$\omega^n Y = Ay$$

so

$$\omega = A^{(1/n)}$$

And the solution to the equation is

$$y = y_0 \exp\left(A^{(1/n)}t\right)$$

If A was positive, or n was odd, we're done. If A was negative, and n was even, then ω is imaginary. This isn't a problem because $e^{iAt} = \cos(At) + i\sin(At)$, so an imaginary ω just means that you have oscillatory behavior instead of exponential. Fortunately, basically everything in this class is cosine/sine or exponential, so this is probably the only technique you need. Also, I solved the most general case here, so there isn't much left to do.

2.7 A derivation (optional)

The following derivation to prove the statement that to increase a function the most with a fixed step we should increase y also, even though the largest partial is in x. This might be enlightening, but also tricky and not strictly necessary for understanding, so feel free to skip or skim it if you want.

Note that our total step distance must $\Delta r = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$ must be fixed. Because the step size is so dominated by the largest value, even if there is a relatively small rate of change in some direction it is still beneficial to move a bit in that direction. For a concrete example, imagine that we were considering moving only in Δx because that provides the steepest change. That means $(\Delta y)^2 = 0$, but

‖ Is it obvious that I hate differential equations?

quadratic functions grow slower than linear functions for very small values, so we can add more to Δy than we needed to take away from Δx and keep the step size the same.

Formally (using that $\Delta z = 0$ for clarity)

$$\Delta r = \Delta x \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2}$$

If we subtract some ϵ from Δx, we can add some unknown β to Δy and keep Δr constant, so we set them equal

$$\Delta x \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2} = \Delta x \sqrt{1 - \epsilon + (1+\beta)^2 \left(\frac{\Delta y}{\Delta x}\right)^2}$$

Canceling Δx from both sides and taylor expanding both sides in terms of $z = \frac{\Delta y}{\Delta x}$ around $z = 0$

$$1 + \frac{z^2}{2} = \sqrt{1-\epsilon} + \frac{(1+\beta)^2 z^2}{2\sqrt{1-\epsilon}}$$

Taylor expanding again, but this time for ϵ, (and dropping higher order terms)

$$1 + \frac{z^2}{2} = 1 - \frac{\epsilon}{2} + \frac{(1+\beta)^2 z^2}{2}$$

Collecting terms

$$\frac{z^2}{2}\left(1 - (1+\beta)^2\right) = -\frac{\epsilon}{2}$$

Taylor expanding for β

$$\frac{z^2}{2}(-2\beta) = -\frac{\epsilon}{2}$$

Solving for β

$$\beta = \frac{\epsilon}{2z^2}$$

but

$$z = \frac{\Delta y}{\Delta x}$$

which was almost 0. So for very small values, the amount we add to Δy by subtracting a small amount from Δx is HUGE (formally trends to ∞). Thus, as long as our function varies at all in y, we will want to step at least a small amount in that direction.

Part II

Mechanics

Chapter 3

Gravity

3.1 Basics of gravity

You have all been studying gravity for years, so I won't waste much time here. You already know that the magnitude of the gravitational force between two objects is given by

$$F = \frac{GMm}{r^2}$$

We are concerned with fields in this class, so we will instead define the gravitational field vector. This is a vector that we can multiply by any test mass to get the force on that mass. It is

$$\vec{g}(\vec{r}) = -\frac{GMm}{r^2}\hat{\mathrm{r}} \tag{3.1}$$

Where $\hat{\mathrm{r}}$ is the vector that points from the mass we are calculating the field of to the point where the field is being calculated. The negative sign indicates that the test particle will fall toward the mass rather than being pushed away. The notation $\vec{g}(\vec{r})$ probably looks unfamiliar to you. I could have written $\vec{g}(x, y, z)$, but this would only make sense in Cartesian coordinates. In spherical the equivalent would be $\vec{g}(r, \theta, \phi)$. Since we don't know what coordinate system we will eventually be using, we use $\vec{r}$ in order to be as general as possible.

It is very important to note here, that the r and $\hat{r}$ here do **NOT** correspond to the spherical unit vectors. Instead they represent the magnitude and direction of a vector that points from the mass to the point we are calculating at.* This distinction will be important sometimes for gravity, but start being careful now, because in electrostatics, it will be tricky to not mess this up.

Now we will make use of something called the superposition principle, which is a cornerstone of physics. Essentially, the superposition principle says that

$$f(a + b + c) = f(a) + f(b) + f(c)$$

provided f is linear. In words: **the solution of a linear equation for a sum of inputs is the sum of the solutions for each input**. This probably sounds obvious, but it gives us a less than obvious result. If we want the field of a bunch of masses, we can find the fields from each mass and add them up.†
You have, of course, been using the superposition principle constantly without knowing the name. If I asked

*If you are wondering why we can't just take the mass to be at the origin, the multiple mass case makes it clearer. Which mass would we place at the origin? We can only pick one...

†If you object that this is a vector equation, and it isn't linear, you are correct, but it only needs to be linear in the mass, and we can treat each direction separately.

you to find the kinetic energy of a system, you would know to add up all the individual kinetic energies.

Using superposition we can say that the gravity field from set of masses should be

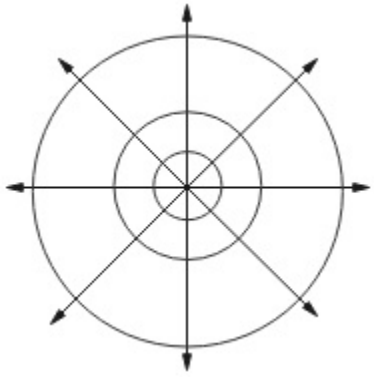
$$\vec{g}(\vec{r}) = \sum_i \vec{g}_i(\vec{r}) = \sum_i -\frac{GM_i}{r_i^2}\hat{\mathrm{r}}_i \tag{3.2}$$

The vector notation here makes it explicit that the result does not depend on coordinate system. That said, evaluating the sum in a system that isn't Cartesian is unlikely to be productive. It is the fact that the unit vectors are constant everywhere that makes Cartesian coordinates a good choice here.

3.2 Gauss' Law

So far, nothing was really new except maybe the notation. Now, however, we are going to focus on the $\frac{1}{r^2}$ in the denominator. For this we are going to think about a completely different system. Consider a lightbulb that radiates evenly in every direction. The light is producing photons, which will travel in straight lines. Only a tiny fraction are absorbed by air, so its actually quite a good approximation to say that all the light eventually hits a solid surface. Take a second and think about why the light is getting dimmer as you walk further away and if you can, try to derive the dependence of number of photons hitting your eye each second on distance (ignoring constants, of course).[‡]

To analyze the light, start by drawing a picture like the one below. Imagine that a very short pulse is emitted, and the lines are the trajectories of photons. Notice that every line that goes through the innermost circle also goes through the other circles, but that the impact points get further apart on the larger circles than they were on the smaller ones. If you look more carefully, you will notice that the angle between the rays does not change: the number of rays in a given angle is conserved as we move outwards.



I have drawn this in 2D, so let's go ahead and derive the answer in a 2D universe. In that universe, if we call the radius of the innermost circle r_1, then the distance between the rays on that circle is

$$D_1 = \frac{2\pi r_1}{8} = \frac{\pi}{4} r_1$$

For the second circle it would of course be

$$D_2 = \frac{\pi}{4} r_2$$

more generally, we could say that

$$D(r) = \frac{\pi}{4} r$$

[‡]If you try this experiment, your perceived brightness will change with distance, but not the way we derive. That is because your eyes are approximately logarithmic, not linear.

That means that the density of rays should be

$$\lambda(r) = \frac{1}{D(r)} = \frac{4}{\pi r} \propto \frac{1}{r}$$

Our universe, of course is actually $3D$. If we did this in $3D$, we would have spheres instead of circles. The surface area of a sphere is

$$4\pi r^2$$

This means that the area density of light rays should obey

$$\sigma(r) \propto \frac{1}{r^2}$$

In other words, brightness of a bulb should follow a $\frac{1}{r^2}$ law. The trick here is that in order to get this, we used that every ray that traveled through the first sphere also traveled through the last sphere; in other words, photons number is conserved as we move outward. Or equivalently the number of rays in any solid angle is conserved (solid angle is a name for a 2D angle, basically a cone starting on the center).

The interesting part is that this matches the $\frac{1}{r^2}$ falloff of gravity. Working backwards we can thus make a few interesting statements about the gravity field

1. The falloff of gravity is completely due to the geometry of space
2. The amount of gravity field in any cone beginning at the center of the mass is conserved with distance
3. The amount of gravity field that passes through any surface that contains the mass is the same and depends only on the mass (and constants).

If you accepted these statements as readily apparent or as following naturally from the argument, go back and re-read them carefully. As written, they are far more broad then my simple analysis showed, and I have not proved them yet. In particular, there are three corollaries that follow directly from the third statement.

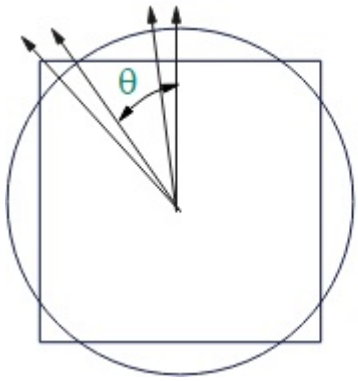
4. The shape of the surface is irrelevant
5. It doesn't matter where inside the circle the massive object is, or what shape it is.
6. The gravity field passing through any surface that does NOT contain mass is 0.

Of the second set of statements, (4) is the easiest to reason. If we go back to the picture and imagine using squares (or cubes in 3D), the total number of rays going through the square would not be different from the circle.
We can do something similar for (5). Look at the picture and imagine moving one of the circles off center. Did that change the total number of rays passing through? Unless you moved the circle enough so that it no longer contained the origin point, the answer is no.
Finally, for (6), displace one of the circles a bit more, so that it no longer contains the origin point of the rays. Notice that now every ray that enters the surface leaves it. If we call entering positive and leaving negative, the net field passing through will give 0.

Now we have some intuition. It's time to formalize it into math. To do this we think back to our mathematical prerequisites chapter, specifically sections 2.5.1 and 2.4.3. Intuitively, if we are looking to quantify the penetration through a surface, we need to think about how the field meets the surface. Consider the example on the next page:

Figure 3.1: The flux through any surface is constant, but the angle that the vectors make with the shape changes.

(a) Flux through a star

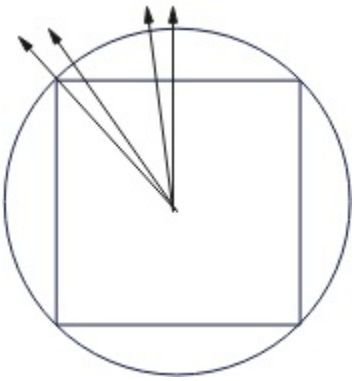


(b) Flux through a square with angle shown

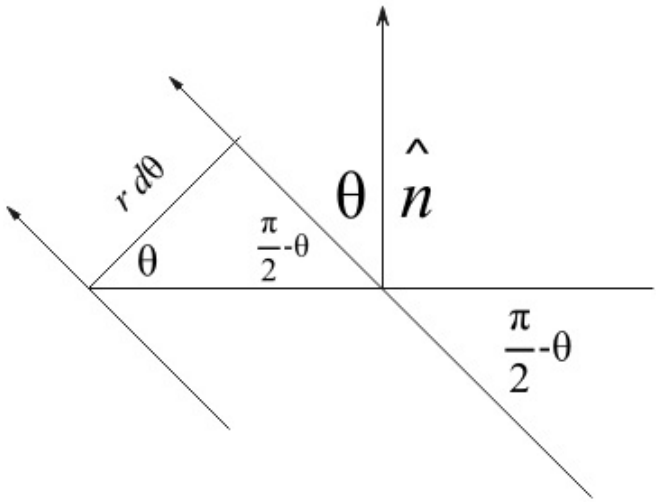
The star clearly has a bigger surface area than the circle does, but notice that the intersections of the rays with the circle are much closer together (measured along the shape) than the intersection of the rays with the star. This is the intuitive explanation for how they can have different surface areas and still have the same total penetration.

Unfortunately, this intuition doesn't immediately give us a math equation. Instead, we will need to look at the angle that the rays make with the shapes. For the circle, they are always exactly perpendicular. For the star, they are very much not. The effect is more visible on the star, but to get the magnitude, we'll need the simple geometry of a square. Consider the two pairs of rays shown in 3.1b. Notice that the square exhibits the same effect as the star.

Figure 3.2: The flux through any surface is constant, but the angle that the vectors make with the shape changes.



(a) Flux through a square



(b) Zooming way in on the angle part and taking $d\theta \to 0$ shows us the geometry clearly

To get further, we will need to resize our square as shown in Each is separated by the same small angle, that we will call, $d\theta$. If we call the center of a side 0 rad, then the further we get from the center, the larger the angle θ between the first ray and the center of a side gets. The distance between rays along the circle is

$rd\theta$ where r is the radius of the circle (from the definition of radian).

Now look at 3.2a. Notice that for the rays near the corner, the distance along the circle is less than the distance along the circle, even though the circle is further away. In the limit that $d\theta$ is infinitesimal, we can treat the "triangle" bounded by the ray pair, the circles edge, and the square as an actual triangle. An enlarged version is shown in fig 3.2b.

In the infinitesimal case, the square side is the hypotenuse. Since the circle side had length $r\ d\theta$, the square side should have length $\frac{r\ d\theta}{\cos\theta}$. We now only have to observe that $\cos\theta$ is also the result of taking $\hat{v}\cdot\hat{n}$, where $\hat{v}$ is direction of the ray, and $\hat{n}$ has its standard definition as the unit outward for the square.

This was a lot of math, but the result is useful. Since we showed it was the same angle, we will now re-define θ as the angle between the vector and $\hat{n}$. This means that we have the result that, as the vector gets further from being perpendicular to the surface, the rays spread out as $\cos\theta$.

Now I will provide a quick summary of the argument put out in the last two pages. Our formula for flux was $\int \vec{v}\cdot\hat{n}dA$. In the circle case, our $\vec{v}$ and $\hat{n}$ were already perpendicular, so that $\vec{v}\cdot\hat{n}$ was just v. When we switched to a square, the dA got bigger by $\frac{1}{\cos\theta}$, but the $\vec{v}\cdot\hat{n}$ got smaller by the same factor, leaving the integral unchanged. I showed this for a square, but our values were infinitesimal, so they don't actually care about global geometry. The result holds for any surface.

We will define the quantity $\vec{v}\cdot\hat{n}(\vec{r})\ dA$ as the flux of the vector field $\vec{v}(r)$ through some infinitesimal piece of the surface. Then the total flux should be

$$\boxed{\Phi_e = \int \vec{v}\cdot\hat{n}\ dA} \tag{3.3}$$

where $\vec{v}$ and $\hat{n}$ are used as shorthand for each vector evaluated at each infinitesimal surface as we integrate.

We can at last write out the whole mathematical statement. Using statement (3) about gravity from earlier in the section, we know another expression for the total flux through the surface. It is just the enclosed mass times some constant. Then we can write

$$\oint_{\partial V} \vec{g}\cdot\hat{n}\ dA = KM_{enc} \tag{3.4}$$

Where the $\oint$ symbols reminds us that we are integrating over a closed surface and the ∂V tells us it is the surface bounding a volume that includes the mass M_{enc}. The constant (K) could be determined either experimentally or by comparing to our known equation for g. We will do it the latter way.

Finding the Constant

Start with equation (3.4)and considering a planet with mass M. We can choose whatever surface we want, but the obvious choice is a sphere with radius r centered around the center of the planet (planets are very nearly spherically symmetric, and we want to choose something that maintains that symmetry).

With this choice, we have two really nice simplifications we can make to $\vec{g}\cdot\hat{n}$:

- Because the sphere is centered around the center of the planet, $\vec{g}$ is always exactly **anti-parallel** to $\hat{n}$. It is anti-parallel because $\hat{n}$ is the unit outward facing normal, and gravity always points towards the center of the circle, not away from it.
- The value of $g(\vec{r})$ is the same at every point on the surface of the sphere.

This means that

$$\oint_{\partial V} \vec{g} \cdot \hat{n} \, dA = -4\pi r^2 g$$

Where the negative sign is from the anti-parallel vector dot product. Putting this simplification into Gauss' law, we get

$$-4\pi r^2 g = KM$$

which simplifies to

$$g = -\frac{KM}{4\pi r^2}$$

We know that

$$g = \frac{GM}{r^2}$$

so

$$K = -4\pi G$$

Which finally gives

$$\boxed{\oint_{\partial V} \vec{g} \cdot \hat{n} \, dA = -4\pi G M_{enc}} \tag{3.5}$$

Gauss' Law is one of the most important formulas in this class, and will be vital to build an intuition for the electric field, so make sure you understand it completely...

3.2.1 Using Gauss' Law

The goal of the last section was the determine the unknown constant, but it served as a pretty good demonstration of how to use Gauss' law. Gauss' law is amazing in situations where the very difficult integral on the left side becomes trivial. The general procedure with some commentary is:

1. Find some symmetry in the problem. Nearly always this is either cylindrical or spherical.
2. Define some surface with the same shape as the symmetry in the problem.
   (a) Remember that your surface can't change the answer, so pick the easiest one.
   (b) Both the field and the object must be symmetric.
   (c) Technically the symmetry must be perfect. A finite disk is NOT symmetric because the field must get weaker away from the center. This means that the field is not the same everywhere on the surface, so we cannot easily evaluate the integral.
   (d) In practice problems will state that an object is very large, this is usually physicist code for treat it as infinite.
3. In general, you might be given a density instead of a mass. This is fine as long as the density function is integrable over the object.

3.2.2 Gauss Law and the Divergence Theorem (Optional)

This subsection will not be directly tested by me or the AP. That said, it might be useful to understand for the purpose of this class, and it **WILL** be necessary to understand for future courses, since this gives the form of the equation that is typically used in practice.

It was briefly discussed that the only way there can be a flux through a surface is if something inside is sourcing it. We quantified this by saying that the integral through the volume of the sources inside is equal to the flux through the surface.

$$\oint_{\partial V} \vec{u} \cdot \hat{n} \; dA = \int_V \vec{\nabla} \cdot \vec{u} \; dV$$

The source term $\vec{\nabla} \cdot \vec{u}$ is nonzero only where mass exists. If we take the case of gravity, we already know how to evaluate the left. It is just $-4\pi G M_{enc}$. In turn we can write M_{enc} as

$$M_{enc} = \int_V \rho(\vec{r}) \; dV$$

Substituting this into our equation above

$$-4\pi G \int_V \rho(\vec{r}) \; dV = \int_V \vec{\nabla} \cdot \vec{g} \; dV$$

This must hold over any volume V. This is only possible if the integrands match exactly, so we have

$$\boxed{\vec{\nabla} \cdot \vec{g} = -4\pi G \rho} \tag{3.6}$$

Obviously ρ and $\vec{g}$ are still functions of $\vec{r}$, but that doesn't matter because this statement is local: it applied at each point separately. Essentially this is a statement that **mass sources the gravity field**. It turns out we can derive all the properties of gravity from this very simple statement (and an experiment to get the constant)[§].

[§]Technically we also need $\vec{\nabla} \times \vec{g} = 0$ or equivalently a statement that potential energy is well defined at a point

Chapter 4

Forces, Momentum, Kinematics

4.1 Calculus Updates

You almost certainly know that $F = ma$*. This isn't, strictly speaking, true. Newton's second law is

$$\boxed{\frac{d\vec{p}}{dt} = \vec{F}} \tag{4.1}$$

Using the definition $\vec{p} = m\vec{v}$ and the chain rule

$$m\frac{d\vec{v}}{dt} + \vec{v}\frac{dm}{dt} = \vec{F}$$

Or

$$F = ma + v\dot{m}$$

The second term is usually not important because it makes sense to think of mass as unchanging for solid objects. But consider a rocket. The rocket's mass is absolutely changing with time because fuel is a substantial fraction of the mass of rockets (typically no less than 80% of the mass of a rocket at liftoff is fuel). So we really can't ignore the second term in all situations. We will do the rocket example eventually, but for most things in this class, we needn't worry to much about the second term.
What of course does change is that we can now deal with non constant forces. Integrating both sides gives

$$\vec{p} = \int \vec{F}\ dt$$

Which is a generalization of the law of impulse to non constant forces. Dividing by m on both sides (assuming it is constant!)

$$\vec{v} = \vec{v}_0 + \int \vec{a}\ dt$$

Where I explicitly pulled out the integration constant. Which is our kinematic equation for non-constant forces integrating again for time and recognizing that $\vec{v} = \dot{\vec{r}}$

$$\vec{r} = \vec{r}_0 + \int \vec{v}\ dt$$

or, being explicit about acceleration

$$\vec{r} = \vec{r}_0 + \int \left(\vec{v}_0 + \int \vec{a}\ dt\right) dt$$

*If you don't, find out if it is too late to drop the class

Notice that if $\vec{a}$ is constant with time, we are right back to $\vec{r} = \vec{r}_0 + \vec{v}_0 t + \frac{1}{2}\vec{a}t^2$ as we should be.
With calculus, the kinematic equations come even more trivially from the momentum equation. Just as before, the physics is energy and momentum (or in this case just momentum), kinematics is just math.

That said, we can and should drop our crutch of thinking about integration as area under the curve. A better way of thinking about it is that if we add up the speed at each time multiplied by the time we had that speed, we get how far we went. No reference to area under the curve was ever needed here.

4.2 Drag Forces

With drag forces we can actually do something completely new. Without calculus, we could only talk about the drag force at a moment in time, or sketch the curve based on intuition. Now we can actually do the math. First though, lets go through the intuition.

In the absence of velocity dependent forces, objects will keep the same acceleration forever. When a force is added that increases in magnitude with increasing speed, the object will now reach some equilibrium once the speed dependent force gets big enough to counter the speed independent one.
Drag forces come about from two mechanisms, and they produce slightly different equations (with way different solutions).

4.2.1 Non-viscous Drag

Water is an example of a liquid that, is non viscous.[†] Consider and object moving through the air with velocity $v\hat{z}$ as show in fig. When it hits an air molecule, it will make that molecule move with some speed that is proportional to v.

The constant of proportionality[‡] will depend on the shape of the object. For a sharply pointed object, the particles will barely deflect, so the constant will be small. For a completely flat object, the particles will be pushed along at basically v, so the constant will be almost 1.

If we consider the number of particles hit as we move some short time t, we should hit $Anvt$ particles where n is the number density (particles per unit volume). Each particle will transfer the drag coefficient C multiplied by its change in momentum. If the mass of a molecule is μ, its momentum change would be about μv the rate of momentum transfer out of the moving object by collision should be

$$\frac{d\vec{p}}{dt} \approx -C\mu v n v A \hat{v}$$

Note that when v is 0, the collisions transfer no momentum. This makes sense because in the objects frame, the particles are static on average when the object isn't moving. Later as we speed up, they will have a large momentum upward in the object's frame. Still, the problem is easier to work with in the static frame.

To get a differential equation, we have to put this in terms either entirely of $\vec{v}$ or entirely of $\vec{p}$. We also can recognize that the number of particles per unit volume times the mass of a particle is just the standard mass density ρ

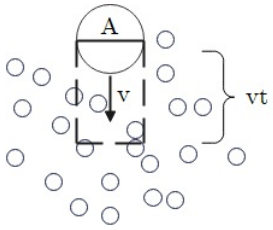
$$m\dot{v} \approx -CA\rho v^2$$

In practice, everyone pulls a factor of $\frac{1}{2}$ out of the C we came up with. So in terms of their C.

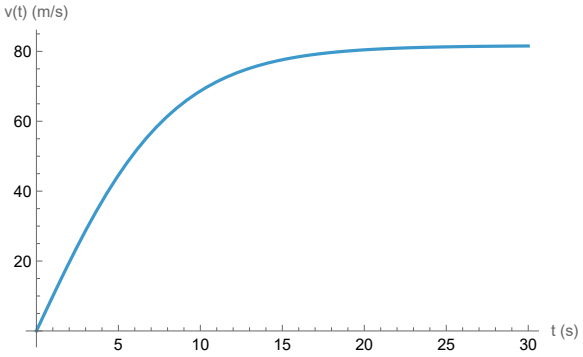
[†] Actually, the importance of viscosity depends on the size and speed of objects. Very small objects moving slowly are highly affected by viscosity. Cellular biology would be very different if water actually had no viscosity.

[‡] It's not actually constant, it varies with speed, size, and viscosity, but typically it varies so slowly that outside of very sensitive cases like aviation, no one cares much

Figure 4.1: A sphere falls through air



(a) As the sphere moves for some very short time t, it will run into all the particles in the cylinder outlined by dashed lines.



(b) The velocity as a function of time will look like this. For a sphere with mass 1kg and cross sectional area 0.03 m^2 falling through air, we see that we hit terminal speed of about 80 $\frac{m}{s}$ and acceleration is almost linear until about 10 s

$$m\dot{v} \approx -\frac{1}{2}\rho v^2 AC$$

Now we just use that there is also a constant rate of momentum transfer into the object from gravity to get the full equation

$$\dot{v} \approx -\frac{1}{2m}\rho v^2 AC + g$$

This differential equation isn't super easy to solve, but Mathematica gives (assuming that we start with speed 0)

$$v(t) = \sqrt{\frac{2gm}{AC\rho}} \tanh\left(t\sqrt{\frac{ACg\rho}{2m}}\right)$$

where tanh is the hyperbolic tangent function (ie tan with an imaginary argument).

4.2.2 Viscous Drag

For viscous drag, we aren't concerned with collisions as much as with stickiness. This means that the above analysis should change because we no longer really care about the momentum of change of the fluid particles. Now the amount of momentum lost to the fluid particles as the sphere interacts should be some constant, nearly independent of the speed of the sphere.
This should immediately tell us that the momentum transfer out of the sphere by drag should look like

$$\frac{d\vec{p}}{dt} \propto \vec{v}$$

We still have the one factor of $\vec{v}$ because moving faster will still results in having more interactions with particles. There really isn't a lot more we can intuit here without more knowledge than we have.
If we wanted to get the answer, the way forward would be dimensional analysis, but I won't do that because you already know how and it provides little insight here.

The answer for a sphere of radius R falling through a viscous fluid is

$$m\frac{d\vec{v}}{dt} = -6\pi\eta R\vec{v} + m\vec{g}$$

or

$$\dot{v} = \frac{-6\pi\eta R v}{m} + g$$

It should strike you as strange that this doesn't depend on the density. Fortunately, it actually does because η is density dependent. The linear dependence on the radius is hard to explain without actually deriving the expression.

This is much easier to integrate.

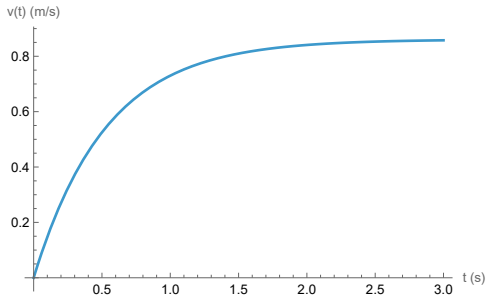
$$v(t) = \frac{gm}{6\pi R\eta}\left(1 - \exp\left(-t\frac{6\pi R\eta}{m}\right)\right)$$

Note that I ignored buoyant force to make the equation a bit simpler. The object would need to be very dense for this approximation to be really accurate. With buoyant force included, we just get an extra upward force based on the volume of the object and the density of the displaced fluid

$$v(t) = \frac{(3m - 4\pi R^3\rho_f)g}{18\pi R\eta}\left(1 - \exp\left(-t\frac{6\pi R\eta}{m}\right)\right)$$

The sphere above would float, so upping the mass of the object to 5kg produces this for the plot. Any real fluid would have both types of drag, but that problem is way too mathematical to be at all enlightening.

Figure 4.2: A sphere falls through honey. Note that we approach terminal speed rapidly, and that the terminal speed is slow.



Chapter 5

Energy

5.1 Basic Formulas

I will emphasize one more time that energy is a mathematical construction. It is extremely useful, but it does not describe a mechanism and it doesn't have a really good definition at this level. That said, we can summarize the formulas we already know

Kinetic energy:

$$K = \frac{1}{2}mv^2$$

Spring Potential:

$$U_s = \frac{1}{k}x^2$$

Gravitational Potential:

$$U_g = -\frac{GM}{R}$$

Rest mass energy

$$E = mc^2$$

Its not immediately obvious what besides units these have in common, but that won't concern us because what we care about is energy conservation.

$$\Delta E = 0$$

For a review of everything you should already know, see the non-calculus based notes. We will continue from there.

5.2 Fields and Path Dependence

Clearly force and energy are related. When we think of a system that stores potential energy, some force is at work there. For most systems, this is ultimately gravity or the electric force, but we won't be too concerned with that. We won't really think about why springs provide a force, and just accept that they do.

To understand this relation between force and energy, consider a mountainous region. If an object ascend the mountain, the system that consists of the object and Earth will have more energy. Something had to put the gravitational energy in. Maybe it was you carrying it. Whatever the case, we could say that the work done was the change in potential energy.

Gravity is a particularly easy case because gravity has the property that it is path independent. It doesn't matter what path you took to carry the object up the mountain. The gravitational energy is the same. You can go in circles all day without ever changing the gravitational potential energy as long as you end at the same point.

This isn't the case for something like friction. If you ride a bike in circles for a long time, friction eventually stops you. Friction wasn't too bad because it always opposed motion, this property meant that the work done was a function only of distance traveled, which was still easy.

For the discussions below, take the system to consist of the object only.

Since we are concerned with fields in this class, the field for gravity from a single source would look something like 5.1b. The field represents the direction of the force at that point. Since power is $P = \vec{F} \cdot \vec{v}$, energy is being added when your velocity has some component along the lines and is being taken away if your velocity has some component anti-parallel to the lines. Try tracing some particle paths with your finger (or print this and use a pencil) through the field in 5.1b. Notice that if you move towards the center, positive work is done on you. If you move away, negative work is done. If you move in a circle with constant distance from the center, no work is done.

Try to find a path that results in you being closer to the center without having a net positive work done. You can't. You also can't find a path that ends the same distance from the center and doesn't have 0 net work done. We would say that the work is path independent here since it depends only on start and end points. Forces that obey this rule are often called conservative forces.

Now we consider a general force. This force $F = F(x, y, z)$ might have the property that the path to get somewhere does matter. Maybe the field will look something like 5.1a. If you look on the left side in the center, you will notice that there is a region where the force is only vertical. This is true almost all the way across the page. This means we can get through the field from the center of one side to the center of the other with almost no work being done.

Now try again, but immediately go down before going across, then go back up when you get to the other side. You are opposing large forces most of the time. This means lots of negative work is done. If you go up first, then across, then down there are now force components that align with you almost everywhere, so a large positive work is done.

This means that we started at the same point, and ended at the same points, but the work was positive, negative, or 0 depending on our route. This all seems quite strange. To make it concrete, imagine water with some whirlpools. In this case, there are forces that act around closed loops. We can actually see one of these in our field. If you look at the whirlpool on the left side of the page, it is possible to go in circles around it while having positive power added forever!

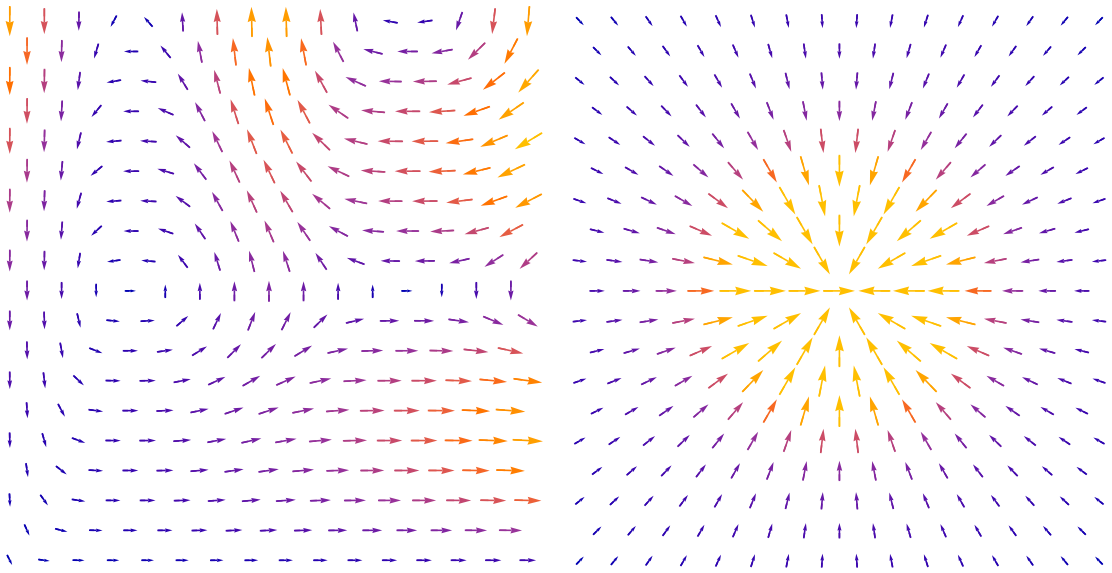
If this seems like it breaks energy conservation, the reason that it doesn't in the whirlpool case is because you cannot actually exceed the speed of the water in the whirlpool. In other cases, it actually is possible to use fields like this to accelerate objects to very high speeds. The trick is that energy must be added to the field to do this, so energy conservation is fine with it.

Hopefully we have an intuition for our goal now. We want to add up work over very small regions to get the total work done, and we need to do this along whatever curve the particle takes. Mathematically

$$\boxed{W = \int_C \vec{F} \cdot d\vec{r}} \tag{5.1}$$

This is the standard definition of work used in most textbooks. I find this confusing to think about, so

Figure 5.1: The size of the arrows is proportional to field strength. If you are looking at this on a computer, the red and yellow are also higher field strength.



(a) Field with large curl. Work is path dependent.

(b) Field with no curl. Work is path independent.

I'll rewrite this

$$W = \int_{t_0}^{t_f} \vec{F} \cdot \frac{d\vec{r}}{dt} dt$$

Which of course is

$$\boxed{W = \int_{t_0}^{t_f} \vec{F} \cdot \vec{v}\ dt} \tag{5.2}$$

This equation fits perfectly with the intuition we went through earlier. To get the total work done we integrate the power that is being added for time.

Whichever you choose, these integrals are not easy. For eq 5.1, you need to remember that $\vec{F}$ is a function of $\vec{r}$, and that the dot product is far from trivial in most cases. If you prefer eq 5.2, the tricky part is that $\vec{F}$ and $\vec{v}$ are both functions of time, but if the functions are known explicitly, this isn't usually as much of an issue.

I did this entire derivation in the context of fields, but that is largely irrelevant. As long as we can get an expression for the infinitesimal work at every point on the path (or power at every point in time), it doesn't matter what produces these forces.

Note that if the integral is path independent we can dodge all this nonsense by just picking a path that is always either parallel, anti-parallel, or perpendicular to the force, so that the integral becomes trivial. In practice, work integrals are rarely useful for mechanics problems where we can't make the integral trivial. Instead, momentum, kinematics (or more often after this class, variational methods with energy) are likely to be used.

If the integral is path independent, then it makes sense to define the potential energy at a point. This will make more sense if we consider one object to be very large and the other very small. Then we can think of potential energy as the work that would need to be done on the small object in order to move from a

reference point to the point of interest. Note that this is exactly the negative of the work that would be done on the small object by the large object as it moved. So

$$U = -W = -\int \vec{F} \cdot d\vec{l} \tag{5.3}$$

Often, the reference point is taken to be infinity so that we get the absolute potential energy. For most purposes though, only changes in energy matter, so other reference points might be used.* Note that for gravity taking our starting point at infinity means that, the potential energy of any point in negative. This is because no energy is ever needed to get to any point closer than infinity, the larger mass will pull us there for free, doing positive work as it does. Equivalently, energy would need to be added to get back to infinity, which is where we defined our reference point.

5.2.1 Evaluating the Integral: sign conventions

The signs for line integrals of potential energy really confused me the first time I saw it, largely because no one really shows all the steps.† After seeing it once, I understood the signs and never wanted to do it explicitly again. Still, I will do it here in the hopes that it clears up confusion.

We start with the equation (5.3)

$$U = -\int \vec{F} \cdot d\vec{l}$$

for gravity

$$U = -\int \vec{F}_g \cdot d\vec{l}$$

Now we plug in for the gravity force, noting that it always points toward the origin (we are taking the larger object to be at the origin and un-moving).

$$U = -\int \frac{GMm}{r^2}(-\hat{\mathrm{r}}) \cdot d\vec{l}$$

Now we imagine taking the charge in from infinity along a straight line. We are always traveling along $-\hat{\mathrm{r}}$, so $d\vec{l} = -\hat{\mathrm{r}}\ dl$

$$U(R) = -\int_{\infty}^{R} \frac{GMm}{r^2}(-\hat{\mathrm{r}}) \cdot (-\hat{\mathrm{r}}) dl$$

recognizing that $-\hat{\mathrm{r}} \cdot (-\hat{\mathrm{r}}) = 1$

$$U(R) = -\int_{\infty}^{R} \frac{GMm}{r^2} dl$$

Its tempting to evaluate this immediately, but we'll get the wrong answer. Instead, recognize that our bounds are wrong for us to integrate along r. Integrating along r would give us the opposite bounds (equivalently $dl = -dr$ because r increases outwards). This means we pick up one more negative sign if we want to keep our current bounds‡ We cancel the new negative sign we got switching to r with the one already there

$$U(R) = \int_{\infty}^{R} \frac{GMm}{r^2} dr$$

The integral is easy now (recognizing that $\frac{1}{\infty} = 0$)

$$U(R) = -\frac{GMm}{R}$$

*We will also see that there are cases where infinity is not valid as a starting point, so other points must be chosen

†Also, there are wrong explanations of the signs EVERYWHERE it seems

‡Incidentally, this is the sign that I didn't see when I learned.

5.2.2 Force From Energy

If an objects potential energy as a function of position is known, it can be used to calculate the force on an object. Obviously, this only works if the integral is path independent, otherwise the idea of energy as a function of positions isn't well posed. To do this note that we did an integral of the force to get energy, so we should take a derivative of energy to get force. The issue is that we did a line integral, and it isn't really clear how to undo that.

Let's start with the 1D case. In 1D, this problem isn't hard because the dot product is trivial: all the information that is relevant is just encoded in the sign anyway. We just use that the force should tend to push us from places of high energy to places of lower energy, so

$$\vec{F}_x = -\frac{dU}{dx}\hat{\mathrm{x}}$$

We could use several different examples here, as we pull a spring from equilibrium, the potential energy increases, and the force pulls us back to equilibrium. As you ascend a mountain, the potential energy increases, and the force tries to pull you back down.

The multi-dimensional case is more interesting. Intuitively, the concept isn't really different. If you are on a mountain and place a soccer ball on the ground, it will tend to roll down the mountain. Now though, we might imagine that the mountain isn't simply a 1D ramp. As the ball goes down, the terrain probably changes, so that it will almost certainly not travel in a straight line as it moves (this is true even ignoring small perturbations from pebbles and such.) Since we don't want to keep changing our coordinate system so that the balls' path stays along one axis, we now have a multi-dimensional problem.

We seek a condition that the ball will follow. Using our prior experiences, things tends to move along whatever direction results in them decreasing their height the fastest. Obviously this would always be straight down, but the ball won't travel through the mountain due to other forces. So it takes the most rapidly decreasing in height while still being allowed. That is our condition. **A conservative force points in the direction of fastest decrease of the potential energy.** Our example convoluted height with potential energy, for gravity they are proportional. For a spring, the direction of fastest decrease in energy is straight back toward equilibrium.

Or, in the language of vector calculus

$$\vec{F} = -\vec{\nabla}U \tag{5.4}$$

This is a remarkable equation because we have three *seemingly* independent components on the left, but only a single independent component on the right. This means that the x,y, and z components of our force aren't really independent. To understand why, look back at diagram 5.1b. Notice that the lines all point straight inwards from a point. If we changed the field so that some of the lines weren't radial, now we would have path dependence again, so we couldn't define U at all. Thus the three components of force aren't independent because that force is constrained to produce a path independent energy. This turns out to be a really strong constraint.

The classic examples of conservative forces are gravity, the spring force, and the electrostatic force. We will study all of these in some detail.

5.2.3 Conservation and Gauss Law

I spent a lot of time in the gravity section talking about the importance of Gauss law and a lot of time here on the idea of conservative forces and path dependence. They are closely related. If the gravity field was able to loop back or disappear, we would not be able to use Gauss law to find the enclosed mass because some of the field would not make it to our sphere. Whatever it is that sources the whirlpool like field from

5.1b, we can be sure Gauss law would not apply. In fact, a field obeying Gauss law actually implies that it is conservative. The reverse is not true, a force that scaled as $\frac{1}{r^3}$ would still be conservative, but it would not obey Gauss law, because the field falls off faster than surface area increases, so the factors don't cancel and the surface used matters. A $\frac{1}{r}$ force has a similar problem, it falls off more slowly than surface area increases, so the flux through a distant surface would be larger than one through a closer surface. If we moved to a 2D universe, our $\frac{1}{r}$ force would obey Gauss law.

Chapter 6

Oscillators

6.1 Spring-Mass

The simplest oscillator we will deal with is the mass on a spring. We already kind of know the answer, but here is a rigorous approach.
We start by noting that $\vec{F} = -k\vec{x}$ for a spring. Then we use $\vec{F} = m\vec{a}$ and $\vec{a} = \ddot{\vec{x}}$. I will drop all the vector signs because we can always just take our axis to be along the spring. Assume the spring is horizontal and the mass is on a table, but it doesn't actually matter. Gravity changes nothing here. We will also assume that we start with the mass pulled all the way to the positive side.

Our equation is

$$m\ddot{x} = -kx$$

Guessing $x = x_0 e^{\omega t}$ we get

$$m\omega^2 = -k$$

or

$$\omega = i\sqrt{\frac{k}{m}}$$

So our solution is

$$x = x_0 \cos \omega t$$

where I discarded the sin solution because we start from equilibrium. We then just get the velocity and acceleration from differentiation

$$v = -x_0 \omega \sin \omega t$$

$$a = -x_0 \omega^2 \cos \omega t$$

That's it. We're done.

6.2 Other Oscillators

Lots of other things oscillate. Many of these act exactly like springs. To solve a general oscillator, we just need to get something in a form that look like $f = -kx$. If we can do that, it's a spring for our purposes.

Consider an object tied to a string and swung around in a circle. The string has some constant tension F and length l the object has mass m. We'll ignore gravity and say motion is in the $x - y$ plane. We want to find the x and y positions as a function of time. We will also assume that at $t = 0$, the string points along

the x axis.

Notice that we can write that the x position when the string makes some angle θ with the x axis is $x = l\cos\theta$ and the force is F pointed toward the center. If we are that angle, the x component of the force is $F_x = -F\cos\theta$. This means that $F_x = -\frac{F}{l}x$. So the x direction behaves like a spring and the motion in x should be solvable using

$$\ddot{x} = -\frac{F}{ml}x$$

Which gives (using that we just solved it in the previous section).

$$x = r\cos t\sqrt{\frac{F}{lm}}$$

We can check easily check this particular case by using a different method: We get the angular speed from

$$\frac{mv^2}{l} = F$$

so $\frac{m\omega^2 l^2}{l} = F$ or $\omega = \sqrt{\frac{F}{lm}}$. Then we use that $x = r\cos\omega t$ and get the same answer.

A relatively easy example was chosen here, but would you have guessed that a spring and circular motion were mathematically the same? The point is that things that seem very not-springlike, may be springlike.

This is quite general and is a classic "clever" solution method. If you can show that it has the spring equation, it can be solved.

Chapter 7

Rotation

7.1 Angular Momentum

It is time to return to everyone's least favorite topic from previous classes: rotation. The change that you all already expect is that we can now properly write the relation between torque and angular momentum

$$\boxed{\frac{d\vec{L}}{dt} = \vec{\tau}} \tag{7.1}$$

Well that wasn't bad actually. Now we need to think about what to do with all these other quantities. Let's start with $\vec{\omega}$. I emphasized that it was a vector, but not how we managed to get a vector out of a change in angle. First, intuitively it has to be a vector, because clearly rotating counterclockwise is distinguishable from clockwise*. Intuitively, if an object is rotating with nothing around, it better be conserving angular momentum. Since angular momentum is a vector, this means the direction can't be changing, so we need to find an appropriate direction. Probably we would first imagine it pointing somewhere in the plane of rotation, this can't work.

There isn't any direction in the plane that is constant and has anything to do with the objects rotation. Then we look at the rotation axis of the object. That is constant, and uniquely defines the objects motion except for the sign. For the sign we use the right hand rule. Curl your fingers along the motion of an actual point on the object. Your thumb will point in the direction of motion.†

With this, we can define $\vec{\omega}$ as the object whose magnitude is $\dot{\theta} = \omega$ and whose direction points according to the right hand rule. It's a good definition, and is what some textbooks use. We'll keep it as our definition. Now, however, we seek an equivalent statement in terms of $\vec{v}$ and $\vec{r}$.

You should be able to quickly convince yourself that the direction of $\vec{r} \times \vec{v}$ is correct by using the right hand rule. Of course the units are wrong and also it should not get bigger as $\vec{r}$ gets bigger. Now recall that $v = \omega r$. With this you might notice that we can just pull out our factors of r from both vectors to get

$$\vec{\omega} = \frac{\vec{r} \times \vec{v}}{r^2} = \hat{r} \times \hat{v} \tag{7.2}$$

This probably looks somewhat bizarre if you are used to thinking in Cartesian coordinates. What does the cross product of two unit vectors have to do with the motion of the object? If you think in terms of magnitude and angle, it makes a lot more sense. Remember than $\vec{r}$ is the vector that points from the origin

*If it wasn't, clocks would be very interesting objects

†There are a ton of different ways to do the right hand rule. This particular one works great for rotation or circulation.

to the object. Then, since we divided out r, $\hat{\mathrm{r}}$ encodes the current angle of the object, while $\hat{v}$ encodes the current angle that the objects velocity if those point in the same direction, the object would be moving either straight toward, or straight away from you, which means no change in angle. If they are perfectly perpendicular, the object is circling around you, which means lots of change in angle.

Got all that? Cool. Now let's try to expand $\frac{d\vec{L}}{dt}$.
Writing out the definition of angular momentum

$$\boxed{\vec{L} = \vec{r} \times \vec{p}} \tag{7.3}$$

$$\tau = m\frac{d(\vec{r} \times \vec{v})}{dt}$$

But we just saw that $\vec{r} \times \vec{v} = r^2\vec{\omega}$

$$\tau = m\frac{d(r^2\vec{\omega})}{dt}$$

which expands to

$$\tau = mr^2\frac{d\vec{\omega}}{dt} + 2mr\vec{\omega}\frac{dr}{dt}$$

For a solid object, the distance a point is on the object is away from the axis isn't changing with pure rotation. In this case we can simplify this to the familiar

$$\tau = I\vec{\alpha}$$

Incidentally another thing we could have done was used that $mr^2 = I$ for a point particle immediately. Then we would have got $\vec{L} = I\vec{\omega}$. Next

$$\frac{d\vec{L}}{dt} = I\frac{d\vec{\omega}}{dt} + \vec{\omega}\frac{dI}{dt}$$

This has the advantage that it looks a lot like the expansion of newton's second law in the linear case. It does make it less obvious why a change in I matters though. If we were getting further from the axis, but maintaining our angular speed, that would certainly not be possible without a torque.
We will deal mostly with solid body rotation, so that second term should usually not be important in calculations, but it is definitely important conceptually.

7.2 Work and Energy

Imagine that we apply a force to an object that can spin around some fixed axis. What should determine the work we can do? If the force is not along the direction of motion at that point, it won't help much, so we want it to be parallel to the linear velocity at the point the force is applied. It would also help if we applied our force far from the center of the disk. Finally, applying it for a longer distance would result in more work done. This all follows from our definition of work. Basically we needed to make $\vec{F} \cdot d\vec{l}$ as big as possible, then do this for as long a distance as possible.
This was all in terms of linear quantities. We seek an answer in terms of angular quantities.

Wanting our force to be close to parallel to infinitesimal displacement is the same as saying that we would like our torque to be close to parallel with the rotation. This again, makes some intuitive sense, applying a large torque around an axis that the object can't rotate around isn't going to be very useful at speeding it up. Applying it against the direction of rotation would just slow it down. So we suspect that the integrand should be $\vec{\tau} \cdot d\vec{\theta}$. So our expression should be

$$\boxed{W = \int_{\theta_i}^{\theta_f} \vec{\tau} \cdot d\vec{\theta}} \tag{7.4}$$

Part III

Electricity and Magnetism

Chapter 8

Electric Fields

8.1 Definition and Intuition

Just like with gravity, we are often interested in finding the electric force. We recall that the $g \propto \frac{1}{r^2}$ term in the gravity force came about because of the geometry of space; we live in a 3-D universe, so the magnitude of the field falls off as $g \propto \frac{1}{r^2}$ because that is the surface area of a sphere with radius r. Obviously, I could have written that it was $g \propto \frac{1}{4\pi r^2}$ instead, but that wouldn't really help much because we know there will be a constant anyway, so we might as well put the 4π into the constant to make the formula nice.

For electrostatics, we will note that we still live in the same universe as before, so it makes sense that the law for electric field (E) should also have the property $E \propto \frac{1}{r^2}$. Indeed it does. If the electric force was always attractive, like gravity, the field of a point charge would always go directly toward the charge as it did with gravity. We know that the electric field can actually be attractive or repulsive, so we conclude that, for a point charge, the field should be given by

$$\vec{E} = \frac{kq}{r^2}\hat{r}$$

where $\hat{r}$ is the unit vector that points directly from the charge producing the field to the point where the field is being calculated. The explicit value of k in SI units is

$$k = \frac{1}{4\pi\epsilon_0} = 8.99 \times 10^9 \ \frac{\mathrm{N \cdot m^2}}{\mathrm{C^2}}$$

ϵ_0 is another constant. Its value is

$$\epsilon_0 = 8.8 \times 10^{-12} \ \frac{\mathrm{C}}{\mathrm{V \cdot m}}$$

Now we once again use the superposition principle. If we want the field from multiple charges, we can just find the field from each charge and sum them. We need to be careful though because the directions of the fields are not in general the same, so we have to add up the vector components. The general formula is

$$\vec{E} = \sum_i \frac{kq_i}{r_i^2}\hat{\mathrm{r}}_i \tag{8.1}$$

Where r_i is the distance from the particle we are currently considering to the point we are calculating the field at and $\hat{\mathrm{r}}_i$ is the unit vector that points from that particle to the calculation point. In practice, you will likely need to put each vector in Cartesian coordinates to evaluate the sum.

8.2 Calculation for a Charge Distribution

I hinted that the superposition principle was going to be useful for us, and it will even more so now that we want to consider extended charge distributions. If we wanted the field of an extended object, we could divide it into a bunch of really small pieces and then sum up the electric field for all these parts using eq 8.1. While this would theoretically work, we already know a better way. For continuous objects, the sum converges to an integral. This gives us the formula

$$\vec{E} = \int \frac{k}{r^2}\hat{\mathrm{r}}\, dq \tag{8.2}$$

Which at first glance doesn't look so bad... But when we think about it, that $\hat{\mathrm{r}}$ term seems like it is going to be a menace. It is. Remember that in the sum $\hat{\mathrm{r}}$ represented the unit vector pointing from the charge to the point of calculation. This didn't change for our integral. It now represents the unit vector that points from our infinitesimal unit of charge to the calculation point. The problem is that this is different for every point! The same issue, of course applies to r, but it is at least a bit more familiar because at least r changing affects only the magnitude, and not the direction.

The most important thing to note here is what r and $\hat{\mathrm{r}}$ are NOT: r **and** $\hat{\mathrm{r}}$ **do NOT represent the spherical or cylindrical unit vectors or coordinates.** Spherical and cylindrical unit vectors start at the origin and point to the charge, we are integrating over something that starts at the charge and points to the calculation point. $\hat{\mathrm{r}}$ will be a menace indeed.

Warning aside, we can write this more explicitly as

$$\vec{E} = \int \rho(\vec{r})\frac{k}{r^2}\hat{\mathrm{r}}\, dV \tag{8.3}$$

Where dV denotes a volume integral and $\rho(\vec{r})$ is the volumetric charge density, which in general is a function of position. Really the volume integral is three separate integrals over whatever your coordinates happen to be. In some cases, we may be given surface (σ) or linear (λ) charge density instead of volumetric. In this case, part of the integral is essentially done for us, so the integrals becomes

$$\vec{E} = \int \sigma(\vec{r})\frac{k}{r^2}\hat{\mathrm{r}}\, dA \tag{8.4}$$

or

$$\vec{E} = \int \lambda(\vec{r})\frac{k}{r^2}\hat{\mathrm{r}}\, dl \tag{8.5}$$

Note that all of this works just as well for gravity as it does for electromagnetism. The only reason it wasn't covered earlier is to give students who are taking their first year of calculus concurrently some breathing room. We will go back and apply this all to gravity at some point.
If you care about a discussion of notation see below, otherwise go ahead and skip it.

8.2.1 A Note on Notation (Optional)

You can skip to the start of the next section if you want, but for some a discussion on the notation used might be enlightening. I want to make a note about notation before moving on. There are many different alternate notations used for equation 8.1 and 8.3, each with its own adherents who swear theirs is the clearest way. I looked through the textbooks I had, and in three books found three very different notations. Jackson, for example, uses

$$\vec{E} = \int \rho(\vec{x}')\frac{\vec{x}-\vec{x}'}{|\vec{x}-\vec{x}'|^3}d^3\vec{x}' \tag{8.6}$$

while Griffith uses

$$\vec{E(\mathbf{r})} = \int \frac{k}{\mathcal{r}^{\,2}} \hat{\boldsymbol{r}} \;\; dq \tag{8.7}$$

Where $\hat{\boldsymbol{r}}$ is defined to point from the charge to the calculation point as desired.
The point to all of this is that there is a trade-off between conciseness of notation and clarity. The equation I gave is the one from the AP. It opts entirely for conciseness. There are some advantages to this choice, but it really makes it easy to forget what you are doing. I don't typically end up using any of the formulas here for this reason. I prefer to just understand what I am doing and write out something that works for the problem in question. The only reason we are having this discussion here, and not in the gravity section, is that the AP does not give an integral equation for gravity. Take that as you will.

8.3 A Simpler Way: Gauss' Law

Just like we did for gravity, we can sometimes exploit symmetry to find the field certain geometries. Go back and read the section on Gauss' law for gravity. In order to get that we just had to assume that field lines couldn't disappear, or equivalently that gravity was a $\frac{1}{r^2}$ force. We just saw that the electrostatic force is also. That means that everything from that chapter still applies. We only need to find the constant and think about the signs since the electrostatic force can be repulsive.

We'll start with (using α for the constant so we don't confuse it with k)

$$\oint_{\partial V} \vec{E} \cdot d\vec{A} = \alpha Q_{enc}$$

We will take a sphere with a charge Q and we will take our surface to be a sphere with radius R centered around the charged object. Then we have

$$4\pi R^2 E = \alpha Q$$

or

$$E = \frac{\alpha Q}{4\pi r^2}$$

Coulomb's law says

$$E = \frac{Q}{4\pi\epsilon_0 r^2}$$

Setting them equal, we find that

$$\alpha = \frac{1}{\epsilon_0}$$

So Gauss' law for the electric field is

$$\boxed{\oint_{\partial V} \vec{E} \cdot d\vec{A} = \frac{Q_{enc}}{\epsilon_0}} \tag{8.8}$$

Note that the sign is opposite gravity. We are assuming by default that Q is positive. If a negative values is inserted, the flux would become negative. If you are wondering why the π terms are in different places for the gravity field and the electric field, I don't have a good answer. There is no physics reason for it, so it must be historical.

8.3.1 Differential Form

Like we did in the gravity case, we can then apply the divergence theorem. As a reminder, the divergence theorem states:

$$\oint_{\partial V} \vec{u} \cdot \hat{n} \, dA = \int_V \vec{\nabla} \cdot \vec{u} \, dV$$

Plugging in $\vec{E}$ and using that we already know how to evaluate the left side

$$\frac{Q}{\epsilon_0} = \int_V \vec{\nabla} \cdot \vec{E} \, dV$$

or

$$\int_V \frac{\rho}{\epsilon_0} dV = \int_V \vec{\nabla} \cdot \vec{E} \, dV$$

The integrands must match everywhere if these integrals are to be equal for any volume.

$$\boxed{\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}} \tag{8.9}$$

Essentially this tells us that charge sources the divergence of the electric field and tells us how much electric field charge sources. Remarkably we can derive everything else about electric fields resulting from static situations from this.* This is one of what are called Maxwell's equations. We will get to the rest later.

*Technically we also need that $\vec{\nabla} \times \vec{E} = 0$

Chapter 9

Electric Potential and Energy

9.1 Electric Potential Energy

We already know how to calculate the change in potential energy of a charge. We could go through all the difficulty that we went through for gravity, or we could just recognize that the force law is the same. If gravity was path independent, so is electrostatics. The integral starts as

$$U = \int_C \vec{F} \cdot \ d\vec{r} \tag{9.1}$$

Path independence means that we can unambiguously take the simplest path, which straight against a field line. This makes the integral to calculate the work done quite simple.

$$\int \frac{kq_1q_2}{r^2}\hat{\mathrm{r}} \cdot (-\hat{\mathrm{r}})dR$$

Now we just specify a starting and stopping point

$$U(R) = -\int_{R_0}^{r} \frac{kq_1q_2}{R^2} \ dR$$

evaluating the integral

$$U(R) = -kq_1q_2\left(\frac{1}{R_0} - \frac{1}{r}\right)$$

As with most potential energy expressions, we have the ability to set the reference point for energy. Typically it makes sense to set it at infinity. In this case $r_0 \to \infty$ so our potential energy for two point charges a distance r apart is

$$\boxed{U = \frac{kq_1q_2}{r}}$$

Notice that it is positive when charges are alike. This makes sense, negative potential energy relative to infinity is a characteristic of bound states. No bound state exists with two like charges.

9.2 Definition of Potential

Electric potential (V) is defined as the electrical potential energy (U) per unit charge at a point in space. **Important note: this is the definition given for electrostatics. In electrodynamics it will not make sense.**
Mathematically

$$U = qV \tag{9.2}$$

By examining our previous equation, we can see that the equation for electrical potential for a point charge should be

$$\boxed{V = \frac{kq}{r}} \tag{9.3}$$

where a positive potential results from a positive charge.*

9.3 Potential and field

Another definition, that is much harder to apply, but is still valid for electrodynamics, uses the integral of the electric field.

$$\boxed{V = -\int_C \vec{E} \cdot d\vec{l}} \tag{9.4}$$

The two definitions are equivalent for all static charge distributions.

Notice that, since $\vec{E}$ is the electric force per unit charge, this equation comes directly from the potential energy equation (9.1) by simply dividing both sides by the test charge and recognizing that the force described there must be pushing against the electric field, hence the negative sign.

Going this way makes perfect sense, but it is also possible to go the other way. For a 1D case, this is rather intuitive.

$$E_x = -\frac{dV}{dx}$$

Once again, there is nothing new here. We did exactly the same thing with energy

$$F_x = -\frac{dU}{dx}$$

so it follows that it should work with potential.
More generally

$$\boxed{\vec{E} = -\vec{\nabla} V} \tag{9.5}$$

This means that if we get the potential somehow, calculating the field is trivial.

9.3.1 Equipotentials and Field Lines

A convenient way to visualize the electric field is to draw either lines that point in the direction of the electric field at each point, or draw surfaces that represent the surfaces of equal potential. Note that electric field lines always point along the direction of fastest decrease (ie opposite the gradient). while movement along an equipotential surface must involve no change in energy. We can immediately say that motion along an equipotential has $\vec{E} \cdot \vec{v} = 0$ which means that they are perpendicular to the gradient, and thus to the field lines.

*If it seems like I am paying way too much attention to negative signs that seem obvious, it is because some negative signs we will see soon are very not obvious

9.4 Equipotentials

I find equipotentials much easier to draw than field lines, and once you have equipotentials, the field lines are really easy. Thus my advice is to draw the equipotentials first, here are the rules for equipotentials.

1. Equipotentials represent surfaces of equal potential
2. Equipotentials are drawn closer together where the field is stronger
3. Equipotentials can never cross each other
4. Equipotentials are symmetric when the charges are symmetric
5. Very close to a single charge, the equipotentials are always circles
6. Very far from a finite charge distribution, the surfaces are **usually** circles.

9.5 Electric Field Lines

1. Electric field lines can only start on positive charges, or at infinity
2. Electric field lines can only end on negative charges, or at infinity
3. Electric field lines point from positive charges to negative charges
4. Electric field lines are drawn closer together where the field is stronger
5. The tangent to the electric field line at a point represents the direction of the force on a positively charged particle placed on that point
6. The electric field lines are NOT the lines that particles will follow

Here are some tips to keep in mind when drawing field lines

1. A symmetric charge distribution will have symmetric field lines.
2. If you know how to draw equipotentials, you can use the fact that the field lines are perpendicular to the equipotentials to help you. Equipotentials are usually easier to draw, so this is an advantage.
3. Electric field lines can never cross each other

9.6 Potential For a Distribution

At this point it should be no surprise that the potential for a charge distribution can be calculated by superposition. Energy and the electric field both did, so it should as well. We can pretty easily formulate the integral

$$V(r) = \int \frac{kdq}{r}$$

Now we see the utility of potential. This integral doesn't have any vectors, any cosines to get components, or any dot products. That means it will in general be way easier to evaluate. It's so much easier that in practice if you want the field a viable method is to find the potential and then take the derivative to get the field.

There is a single "gotcha" with the potential. It may not be well defined at all. In the case of distributions that extend to infinity, the potential is usually not uniquely definable. For some cases (like an infinite cylinder that we will talk about in class) we can rescue the potential by defining some arbitrary place to set it to 0. This avoids the fact that the integral diverges. In other cases, like an infinite plane, the potential just isn't a useful concept and we just use the field instead, or talk only about differences between nearby points rather than trying to treat it globally.

9.7 Energy To Assemble a Set of Charges

It probably seems like if we already have the potential, getting the potential energy of the distribution should be trivial. It decidedly isn't. Getting the potential only required us to consider each charge and the potential it created. Getting the potential energy requires us to consider the interactions between each charge with every other charge. That won't be so easy.
Lets say you want to build some macroscopic charged objects out of pieces. You would essentially need to bring all the pieces in from infinitely far away to their current locations. The first piece is free. After that, you would need to pay an increasingly large energy cost to keep adding pieces. For a finite set of charges, there are multiple ways to proceed. I think the easiest method goes like this

1. Put in the first charge
2. Calculate the potential at the point where the next charge goes
3. Put in the next charge.
4. Repeat item 2-3 until finished

Note that the order we bring things in can't matter because the force involved are conservative. This is easy from an algorithmic standpoint, but would take a very long time to do for a large distribution. A shortcut is to realize that we don't actually care how we built it up, only what it looks like now. So for each charge, we need to find the energy associated with each other charge. The energy between charges j and k should be

$$U_{jk} = \frac{kq_j q_k}{r_{jk}}$$

You might want to sum this over j and k, but we have to be slightly careful, doing that would have two issues

- The j=k case would yield ∞ because $r_{jk} = 0$
- We would be counting every interaction twice because for example $(j,k) = (1,2)$ and $(j,k) = (2,1)$ refer to the same potential energy

This is easily fixed by just adjusting the bounds on the sum. We only want to sum over combinations we haven't see yet. This means that we want the value of k to be greater than the value of j (or the reverse, it doesn't matter) If there are N charges

$$U = \sum_{j=1}^{N} \sum_{k=j+1}^{N} \frac{kq_j q_k}{r_{jk}}$$

So the finite case wasn't too bad. This obviously falls apart for the continuous case. For this we use a trick.

- First, calculate the potential of the final distribution.
- Add a single element of charge dq and find the energy. Add this energy to a running total.

- Remove the first dq, returning the potential to the calculated value and add a second dq at a slightly different location
- Repeat steps 2-3 until we have brought in all charge.

This sounds super weird, but it has a huge advantage: we don't need to think about how the charge already added changes the potential as we integrate. This approach gives exactly twice the potential energy because, on average, only half the charge would have been there when we brought the charge in if we were actually forming the distribution. This leads to the integral

$$W = \frac{1}{2}\int V(\vec{r})dq$$

or, using τ for volume to avoid confusion and dropping explicit $\vec{r}$ dependence from ρ and V (it's still there, of course)

$$W = \frac{1}{2}\int V\rho \; d\tau$$

In practice, the integral is much easier to do in terms of the field

$$\boxed{W = \frac{\epsilon_0}{2}\int_{\text{all space}} E^2 d\tau} \tag{9.6}$$

Where E is the magnitude of the electric field... not energy!
I wish I had a way to get this intuitively, but I don't, it comes from a clever use of some vector calculus identities combined with Maxwell's equations. What I can say is that the E^2 dependence makes sense. One of the E came from the potential, and the other came from ρ (or equivalently, we are concerned with pair interactions, so we count everything twice). Don't forget that the integral here is over all space, not just the region with charge. This expression might make you wonder where the potential energy is actually stored. In previous classes I have dodged the question by saying it is stored in the interaction. In this class, it will be best to think of it as being stored in the fields.

One very important thing to note is that the potential energy is quadratic in the field, not linear. **This means that the law of superposition does not apply to energy of charge distributions.** You cannot in general add the potential energies of multiple systems together. You could also have seen this earlier by noting that we were multiplying all pairs together. In this pairwise sense, the law of superposition still works.

9.8 Conductors

Put very simply, conductors allow motion of charged particles, and insulators don't. No real world object is a perfect insulator or a perfect conductor. With enough voltage, anything will conduct. A everyday example is shocking yourself just before you touch a metal object, like a doorknob. The voltages involved are quite high. At those voltages, even air becomes conductive. A more extreme example of this is called an 'arc flash' and happens when objects are put too close to very high voltage power lines.

Most of the conductivity we will talk about has the electrons moving, while the nuclei stay in place. Consequentially, metals are good conductors, other solids are typically not. This is because metals have electrons that are shared between neighboring atoms.[†] In most other materials, moving electrons requires giving the electrons enough energy to bump them into higher states that allow motion. This means very high voltages are required to make these materials conduct. Gasses under normal conditions are always poor conductors.

[†]Chemists call this a 'sea of electrons'

Some more interesting cases are water and plasma. Pure water is a relatively strong insulator, but add a bit of dissolved salt and you suddenly have something that is rather conductive. This should not be taken to mean that there are lots of free electrons in water. There definitely are not. Instead, water allows ionic compounds to be broken up into positive and negative ions[‡]. These ions are then free to move, and there movement allows the flow of current. This is called 'ionic conduction'. In the plasma phase, a substantial fraction of the atoms are ionized, this means that there are both free electrons and positive ions. The free electrons dominate in this case due to their lower mass.

9.8.1 Conductor Rules and Explanations

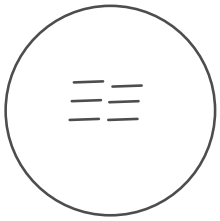
The standard statement of the rules for electric fields inside conductors is as follows:

1. $\vec{E} = 0$ everywhere inside the conductor
2. Just outside the surface of the conductor, $\vec{E}$ is perpendicular to the surface
3. Any net charge on the conductor resides on the surface

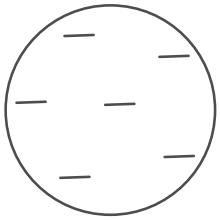
I usually replace the second rule with an equivalent statement that the surface is an equipotential.

No charge inside a conductor

All charge on a conductor must be on the surface. To see why, imagine that we put a bunch of electrons at the center of the conductor, like this:



There will be a strong repulsive force between the electrons that will push them away from each other. That force will be minimized when they are as far apart as they are allowed to move. This happens when they are evenly spread out on the surface of the conductor, like this:



9.8.2 No field inside a conductor

A conductor placed in an electric field will always polarize exactly enough to cancel the external electric field.

To see why, imagine that we place a neutral conductor in an electric field. This will make the electrons move around. They won't stop moving until there is no force on them. This happens when the electric field is 0 everywhere inside the conductor.

[‡]For table salt Na^+ and Cl^-

9.8.3 The surface of a conductor is an equipotential surface

Recall that an equipotential surface is a surface where the electrical potential energy is the same everywhere.

We know that the work done to move a charge between two points is $W = F_{\parallel} D$. This formula says that the work done is the distance that we move against a force times the strength of the force we are moving against. Let's imagine that we move a charge inside the conductor. From the fact that there is no field in the conductor, there is no force on the charge when we move it. This means that any motion inside the conductor involves no change in the energy of the charge. Finally this gives us that anywhere we go on the conductor must have the same potential because we did no work getting there.

This result also implies something closely related: **the electric field on the surface of a conductor must be exactly perpendicular to the surface.** This statement is the more common third rule. It is equivalent to the one given because, if the surface is an equipotential, then all field lines must cross it at right angles, since field lines point along the negative gradient of the field, and equipotentials are perpendicular to the gradient.

Chapter 10

Magnetic Fields

The introduction of magnetic fields marks a critical turn in the focus of these notes. Before this point, I have been using physical examples to try to help build an intuition for the math. The goal was that you would get the math intuitively enough that when we got to something completely new, you could make sense of it in terms of math you already understand. Magnetic fields are that new thing, and they are quite unlike anything in your everyday experience.

10.1 Current and Current Density

You are hopefully already familiar with the idea of current. As a quick refresher, current is the rate that charge flows through some surface (usually this is some section of wire). We denote current I, so we could write

$$\boxed{I = \frac{dQ}{dt}} \tag{10.1}$$

We won't be interested in the current just yet, instead we will focus on the current density. Usually this is denoted $\vec{j}$. Note that we are taking this as a vector. Its direction should point along the direction of motion of the **positive** charges (ie opposite the electron motion). Imagine charges flowing in a wire. We would like to express the current density (eg charge/time/area) in terms of things we already know. We start by noticing that the answer clearly depends on the speed of the charges. If we double the speed, we will double the number of charges crossing any surface. So

$$\vec{j} \propto \vec{v}$$

We also can reason that if the number of charge carriers per unit volume were doubled, this would also double the number crossing

$$\vec{j} \propto n\vec{v}$$

Finally, if we doubled the charge of each carrier, that too would double the amount of charge that crossed

$$\vec{j} \propto nq\vec{v}$$

There isn't any physical reason that this shouldn't just be an equality since no interesting geometry exists here. Indeed, we define it that way

$$\boxed{\vec{j} = nq\vec{v}}$$

Its worth noting here that the product of the number density of charges and the charge of each one nq, is just the charge density ρ. We could have thus equivalently written the current density as

$$\vec{j} = \rho\vec{v}$$

This is what most sources I have seen use for a definition. I also like it better because it makes the analogy to fluids cleaner.
The AP uses $\vec{J}$ instead of $\vec{j}$. This is a strange choice because in physics, lower case letters are typically used to indicate densities, but the choice is up to you.
What if we wanted to get the total current I flowing through a cross section of wire if we knew $\vec{j}$? We actually solved that problem already in section 2.4.3 with water through a hoop. If you need, go review that. Intuitively, we are looking for the flux through a surface. That must depend on the vector field we are integrating ($\vec{j}$) and on the angle that the hoop makes with the field. That was our definition of a surface integral

$$\boxed{I = \int_S \vec{j} \cdot d\vec{A}} \tag{10.2}$$

Typically, we will pretend that current flows uniformly through the wire.*

10.2 Magnetostatics

While the existence of charge creates the electric field, we need charge to be moving to get a magnetic field. We are going to start with an experimental result for this. The result is this: the magnetic field is divergence free. This is among the most tested results in all of science. Despite a century of experiments looking for it, the magnetic field is always found to have exactly 0 divergence. Mathematically

$$\boxed{\vec{\nabla} \cdot \vec{B} = 0} \tag{10.3}$$

or equivalently

$$\boxed{\oint \vec{B} \cdot \ d\vec{A} = 0} \tag{10.4}$$

That is any closed surface we can define has a magnetic flux through it of 0.
If it has no divergence, the only thing it can do is circulate around a point. Where before our task was to understand everything in terms of divergence and surface integrals, we must not use the curl and line integrals. This is going to rely heavily on the information from sections on line integrals (2.4.4) and curl (2.5.1). If you are hazy on how line integrals or curl worked, go back and read those first.

The task from here would be finding out what sources the magnetic field circulation, and how strongly. This can't be done from first principles, and there is no known law to compare to here, so we turn to experiment. The answer is that current sources the magnetic field and that the strength of that sourcing is given by

$$\boxed{\vec{\nabla} \times \vec{B} = \mu_0 \vec{j}} \tag{10.5}$$

where $\vec{j}$ is the current area density.
Using only that current sources magnetic fields, we can already guess what the field of the wire looks like. It can't diverge. If we imagine a long (but finite) straight wire, how do we get it to not diverge? The situation is symmetric.
It can't point radially out from the wire or a Gaussian cylinder around the wire would have nonzero flux.

If it pointed along the wire, it couldn't circulate. This means that it would need to have 0 divergence and 0 curl, which is only possible if it is constant everywhere. To see this, think about a river flowing, if it can't circulate anywhere (ie no whirlpools, and no velocity that is higher on one side than the other since both would allow a closed loop with nonzero line integral), the only option is for it to be totally uniform. If

*This isn't really true, especially at high frequencies, surface effects (called "skin effects") become quite important. That is a topic for another class though

it was uniform, then what would happen at the beginning and end of the wires? We'd need to have it have a nonzero divergence there. It might be possible to skirt all these constraints with some craziness, but a way easier solution is that it circulates in closed loops around the wire.

Now we turn our intuition into math. We can use that the integral of the curl through the disk is equal to the line integral of the field along the boundary (see section 2.5.2). In math, for any field $\vec{F}$

$$\oint_{\partial A} \vec{F} \cdot d\vec{l} = \int_A \vec{\nabla} \times \vec{F} \cdot d\vec{A}$$

Applying that the magnetic field

$$\int_A \vec{\nabla} \times \vec{B} \cdot d\vec{A} = \oint_{\partial A} \vec{B} \cdot d\vec{l}$$

but we know from earlier that $\vec{\nabla} \times \vec{B} = \mu_0 \vec{j}$ so we substitute that in

$$\int_A \mu_0 \vec{j} \cdot d\vec{A} = \oint_{\partial A} \vec{B} \cdot d\vec{l}$$

$\vec{j}$ is the current area density, so integrating over A gives the total current enclosed

$$\boxed{\mu_0 I = \oint \vec{B} \cdot d\vec{l}} \tag{10.6}$$

Both equations (10.5) and (10.6) are referred to as Ampere's law.
So we have another way to look at the field, equivalent to the first. The statement is that currents cause magnetic fields to circulate around wires. We already knew this from just noting that the divergence was 0, so it had to be purely circulatory, but it is good to see this come out of the formal math. Since the magnetic field is a vector, we will need to determine it's direction. For that we will use a right hand rule. If you point the thumb of your right hand along the current, the fingers curl in the direction of the magnetic field.

We now know where magnetic fields come from, but we are still missing important information. What do they do to other objects, and how do we calculate them in situation without enough symmetry to do the line integral?

10.3 Lorentz Force

We have focused up to now on charges that were at rest (or currents in the last section). This is good for a lot of situations, but observant students have probably noticed that things move sometimes. With only static charges, the force on a charged particle is $\vec{F} = q\vec{E}$.
Magnetic fields have a different force law given by $\vec{F} = q\vec{v} \times \vec{B}$. We can combine these to get the Lorentz force law

$$\boxed{\vec{F} = q\left(\vec{E} + \vec{v} \times \vec{B}\right)} \tag{10.7}$$

The cross product with velocity is really interesting. Recall that the cross product takes two vectors and produces a vector perpendicular to both. That means that $\vec{A} \cdot \vec{A} \times \vec{G} = 0$ for any vectors $\vec{A}$ and $\vec{G}$. †
Recall that power is $P = \vec{F} \cdot \vec{v}$
We can immediately say that

$$P = \vec{F} \cdot \vec{v} \times \vec{B} = 0$$

†If you are wondering why there are no parenthesis in some of the formulas, it is because the order is unambiguous, the cross product can accept only vectors.

or more concisely

$$P_{mag} = 0 \tag{10.8}$$

or equivalently the work done is 0 always

$$W_{mag} = 0 \tag{10.9}$$

Notice that I didn't assume anything about the situation here, the Lorentz force law is sometimes taken to be the definition of the $\vec{B}$ field. In any case it is fully general. That means the result is fully general. **Magnetic forces do NO work, EVER.** I am bolding this because there will definitely be times when it looks like the magnetic force did work, but I promise, it didn't (or you will win a Nobel Prize for your discovery).

10.4 Practical Calculation

We seek something equivalent to Coulomb's law, but for magnetic fields. Deriving it from Ampere's law is possible, but is hard enough that it was discovered empirically first. Let's collect what might help and what it tells us

- Magnetic fields have no divergence: we better get something with no radial component
- Magnetic fields are sourced by currents: we better get an integral over the current along the wire
- Magnetic fields curl around current carrying wires: we better get this behavior

We can easily meet the no radial component and curling around the wire parts by having a cross product $d\vec{l} \times \hat{\mathrm{r}}$ where $d\vec{l}$ is an infinitesimal line element of the wire. There are two ways to take the cross product, with opposite signs. I picked here the one that agrees with the right hand rule we discussed earlier. If I had forgotten, we could have used the right hand rule later to figure it out anyway.
To go from a line element to a current, we just multiply by I. This also fits with our intuition that it should be proportional to I.
Additionally, the field of a long wire better fall off with distance from the wire. Since we would be integrating in one spatial dimension, this means that the falloff in the equation should be at least $\frac{1}{r^2}$ for the infinitesimal element. It might fall off faster than $\frac{1}{r^2}$, but that would be strange, since it would imply that the field falling off faster than geometrically. (A stronger argument uses relativity and the fact that the magnetic and electric fields mix at high speeds.) We conclude that the integral better have $\frac{1}{r^2}$
To recap we now have

$$\vec{B} = KI \int_C \frac{d\vec{l} \times \hat{\mathrm{r}}}{r^2}$$

where K is some unknown constant with units. We expect μ_0 to be involved since it was in ampere's law. So we do some dimensional analysis to find (exercise left to the reader, we could also have got the $\frac{1}{r^2}$ dependence here, just remember that differential lengths are still lengths)

$$\vec{B} = D\mu_0 I \int_C \frac{d\vec{l} \times \hat{\mathrm{r}}}{r^2}$$

There is no way to get the dimensionless constant D from these arguments, so we would leave it up to experiment.

$$\boxed{\vec{B} = \frac{\mu_0 I}{4\pi} \int_C \frac{d\vec{l} \times \hat{\mathrm{r}}}{r^2}} \tag{10.10}$$

Which is called the Biot-Savart law.

Taking a look at it makes it clear that this is going to be a mess to evaluate. It's an integral, of a cross product, of two things that both depend on location. Unless the geometry is REALLY nice, it's a problem for a computer. Still, the integral form of Ampere's law was literally unusable unless there was symmetry, and in higher level physics, the problem is solved when you arrive at something that can be handed to a computer. Your job is to formulate the equations, not solve them.

10.5 Force on a Current Carrying Wire

We know that current involves motion of charges, and that magnetic fields cause moving charges to experience a force. It thus seems reasonable to guess that current carrying wires experience a force when placed in a magnetic field. It would be quite difficult to sum this effect up for every charge in the wire, but we can instead use that the current density (ie current per unit area) can be written as

$$\vec{j} = \rho\vec{v}$$

where v is the moving speed of the charges. This means that the current is

$$I = \rho A v$$

We know that the force per unit charge should be

$$\frac{\vec{F}}{Q} = \vec{v} \times \vec{B}$$

Now we notice that we can solve the current equation for v

$$\frac{\vec{F}}{Q} = \frac{I}{\rho A}\hat{v} \times \vec{B}$$

now if we take a tiny segment of wire, with length $d\vec{l}$ the charge inside is

$$dQ = \rho A dl$$

So we can write

$$d\vec{F} = dq\vec{v} \times \vec{B}$$

Substituting our expression for $\vec{v}$, and dQ from earlier

$$d\vec{F} = \rho A \frac{I}{\rho A} dl\hat{v} \times \vec{B} = I dl\hat{v} \times \vec{B}$$

We know that the current flows along the wire, so now we can just orient our dl along the direction of current flow to get $d\vec{l} = dl\hat{v}$, so

$$d\vec{F} = I d\vec{l} \times \vec{B}$$

We can then just integrate to get the total force on the wire.

$$\boxed{\vec{F} = \int I \; d\vec{l} \times \vec{B}} \tag{10.11}$$

Chapter 11

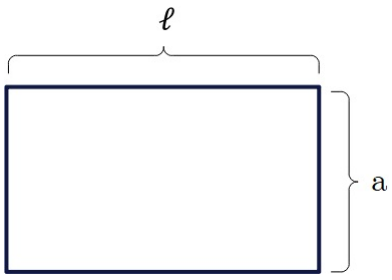
Induction

Consider a rectangular conducting loop, with sides a and ℓ, that is moving through a nonuniform magnetic field with a speed v. We'll take the field to be increasing in the direction of the loop's motion and it's direction to be always exactly perpendicular to the plane of the loop (see diagram). Because the loop is moving, all the charges inside it will have a speed v in the direction of the loops motion. To avoid any more negative signs, we'll pretend that positive charges are moving rather than electrons. Those charges will experience a force

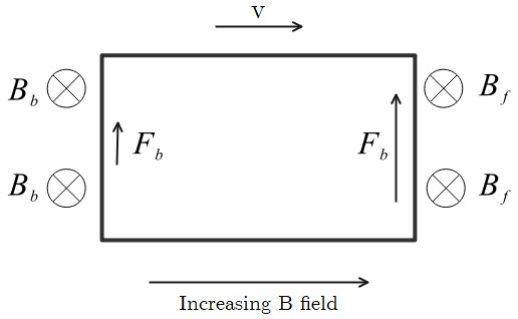
$$\vec{F} = q\vec{v} \times \vec{B}$$

The cross product gives the direction of the force to be upwards for the front and rear bars. For the top and bottom bars, it has no component that points along the bar. Since the field is increasing in the direction of motion, the field will be larger at the front of the loop than at the rear, so the force will be stronger on the charges in the front bar than the rear.

Figure 11.1: A loop travels through a nonuniform magnetic field. A force results that pushes charge around the closed loop.



(a) A loop with sides a and ℓ



(b) A nonuniform field points into the board. The loop moves toward the left, where the field is stronger.

Now let's imagine a charge moving around the loop. We want to find the work done on the charge as it moves through the loop. Lets say that is starts at the bottom left. Then

$$\oint \vec{F} \cdot d\vec{l} = 0 + qvB_f a + 0 - qvB_f a = qva(B_f - B_b)$$

Now we use the tangent line approximation. If we want the value of B at the front, we can find the value at the back and add the change across the loop

$$B_f = B_b + \ell \frac{dB}{dx}$$

We can use this to rewrite our previous expression

$$\oint \vec{F} \cdot d\vec{l} = qva\ell \left(\frac{dB}{dx} \right)$$

Notice that $v = \frac{dx}{dt}$, so multiplication of v with a spatial derivative makes it a time derivative. So we can write

$$\oint \vec{F} \cdot d\vec{l} = qa\ell \frac{dB}{dt}$$

What we found was the work done, but work per unit charge is more helpful since this would give something like voltage. For the moving case, we define this as electromotive force, or EMF and denote it $\mathcal{E}$. With this definition

$$\mathcal{E} = a\ell \frac{dB}{dt}$$

Note that $\frac{dB}{dt}$ was positive, so the fact that I didn't get a negative sign here means that we get a current counter-clockwise around the loop (because that is the direction we did the integral in). Almost everyone writes this law with a negative sign. Physically, that negative sign is meant to serve as a reminder that the effect always opposes the change in flux, but that meaning is easy to lose due to the nonphysical negative sign that results from your choice of a direction of integration. I left it like this rather than reversing integration direction to facilitate this discussion. Fortunately, 11.0.1 will discuss a better way to get directions that relying on negative signs.
Finally, we recognize that

$$a\ell \frac{dB}{dt} = \frac{d\Phi_m}{dt}$$

$$\mathcal{E} = -\frac{d\Phi_m}{dt}$$

We did everything in the frame where we were watching the loop move. We now recognize that if we were in the reference frame of the moving object, the magnetic force couldn't exist because $v = 0$ for the charges. Since this EMF can drive a current through the loop, they better not be allowed to disagree on its value. Since the only other force that could exist is electric, we conclude that in that frame

$$\boxed{\mathcal{E} = \oint \vec{E} \cdot d\vec{l} = -\frac{d\Phi_m}{dt}} \tag{11.1}$$

This relation is called Faraday's Law or sometimes the law of induction. *

I will end with a few thoughts

- I found the work done by the "magnetic force" to be non-zero. This is actually fine because this was just a component of the magnetic force, there is also a component that results from the motion of the charges around the loop. That component does an equal amount of negative work on the system. That said, the whole thing wouldn't have been possible without a magnetic field, so while it didn't do work, it did allow the process to happen.
- The energy put into a charge as it moved around the closed loop was nonzero. This is true and is analogous to our whirlpool example from earlier or to the force of friction when going in circles. In general positive or negative work can be done by forces around closed loops.
- You might have noticed that in addition to nonuniform fields, or fields changing in time, we could also change the magnetic flux by rotating the hoop, or by changing the angles. The formula still works properly in that case.

*After spending a few hours on this section, but not quite being able to prove it, I noticed that Purcell does a similar thing to what I was trying to do in Chapter 7 of Electricity and Magnetism. Unfortunately, his derivation uses a number of things you likely aren't familiar with, but I will acknowledge taking some inspiration from him here.

- I cheated a bit to get to the solution. Technically, the EMF integral should be evaluated at a single point in time (by definition), while the work integral should be evaluated along the physical path that the charge takes, which won't be the same time everywhere. If you want, you can imagine that the time for a charge to move around is short compared to the time for the fields to change substantially.

11.0.1 Direction of the Induced Current

The better way to determine the direction is this physical observation: **the induced current must be in the direction that would create a magnetic field to oppose the change in flux**. This is called Lenz's Law. What this means is that if the magnetic field points upwards, and is increasing, we would need the current to create a magnetic field pointing downward to oppose the change. If the flux was instead decreasing with time, we would need the effect to point upwards to oppose the change. The same logic works also if the magnetic field was in the opposite direction (with the direction switched, of course). I prefer the equivalent statement that the effect acts in a direction that would keep flux as close to constant as possible. A much shorter statement, due to David Griffith is **"Nature abhors a change in flux."** Whichever version you take, the existence of this rule means that we can effectively ignore all the negative sign accounting and recover the direction at the end.

If it isn't obvious why the direction must be this, imagine it was the other way. That would mean that as the strength of a magnetic field increased, we would create a magnetic field that acted to increase the rate of change of the flux. Presumably this increased rate of change would then create a stronger current and so on. This means we could, under the right circumstances, create a self amplifying field with minimal input energy. Obviously this is problematic.

11.1 Mutual Inductance

If we have a set of loops, and we pass a current through one of the loops, there will be a magnetic flux Φ_B through the other loop. This situation happens so often that we define the mutual inductance of two objects

$$\boxed{M_{21} = \frac{\Phi_{21}}{I_1}} \tag{11.2}$$

We would like to get some intuition for what M_{21} could depend on. This is equivalent to asking, with a known current through one, how much magnetic field would pass through the other. There really aren't a lot of choices. We could move the two objects further apart, that would clearly change the answer. We could make one object or the other larger. That would probably change the answer. If the objects are composed of coils, we could also add more coils to one or the other. What all these things have in common is that they are all geometric changes. The mutual inductance is solely a function of the geometry of the arrangement.

What if, instead of the flux induced by 2 by 1, we were interested in the flux induced in 1 by 2. Notice that if we doubled the number of coils in object 1, that would double the flux if it was the reciever, or double the field if it was the initiator. Either way, the answer doubles. If we doubled the distance between them, that would cut the flux by the same factor regardless of whether we were considering the flux through 1 from 2, or 2 from 1. In fact, any change that we can make is symmetric regarding a swap of which coil is creating the field and which is receiving it. This leads to a very helpful result

$$\boxed{M_{21} = M_{12}} \tag{11.3}$$

Note that this does NOT say that we should go through the formula (which might have things like N_2 and R_1) and swap every 1 and 2. Instead, it tells us that if we can calculate M_{21} by any means, we immediately know that M_{12} is identical. This is useful because there will be times where the calculation of M_{21}

is straightforward, but M_{12} would be very difficult from first principles.

By rearranging and taking a derivative, we can immediately notice that

$$\frac{d(M_{21}I_1)}{dt} = \frac{d\Phi_{21}}{dt}$$

We won't be interested in situations where the shape changes, so this gives

$$M_{21}\frac{dI_1}{dt} = \frac{d\Phi_{21}}{dt}$$

or, using Faraday's Law (and reversing LHS and RHS),

$$\mathcal{E}_2 = -M_{21}\frac{dI_1}{dt}$$

Which tells us why we call this mutual inductance. It gives the induced EMF in terms of the rate of change of current.

Chapter 12

Maxwell's Equations

12.1 A Paradox and Its Resolution (Not started yet)

12.2 Maxwell's Equations Together

When we combine everything that we already know, we have the following equation set

$$\vec{\nabla} \cdot \vec{B} = 0 \qquad \vec{\nabla} \times \vec{B} = \mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \tag{12.1}$$

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \qquad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{12.2}$$

or, in integral form

$$\oint_S \vec{B} \cdot d\vec{A} = 0 \qquad \oint_{\partial S} \vec{B} \cdot d\vec{l} = \mu_0 \left(\int_S \vec{j} \cdot d\vec{A} + \epsilon_0 \frac{d}{dt} \int_S \vec{E} \cdot d\vec{A} \right) \tag{12.3}$$

$$\oint_{\partial V} \vec{E} \cdot d\vec{A} = \frac{1}{\epsilon_0} \int_V \rho \, dV \qquad \oint_{\partial S} \vec{E} \cdot d\vec{l} = -\frac{d}{dt} \int_S \vec{B} \cdot d\vec{A} \tag{12.4}$$

These are called Maxwell's equations. They were both the greatest triumph of classical physics, and Einstein's inspiration for relativity.

Note that if there was a such thing as magnetic monopoles (ie magnetic charge), the equations would be symmetric in $\vec{E}$ and $\vec{B}$. This (and some reasons from quantum field theories) have led to a multitude of experiments looking for them. More recently, cosmology has offered a natural answer for their absence: the expansion of the universe would have scattered them so far apart a region of reasonable size is unlikely to have any. Some particle physics theories also predict that magnetic monopoles might would be created in very high energy particle collisions, but would have extremely short lifetimes. Whatever the case, finding the existence of magnetic monopoles would be of theoretical interest, but practically, they would be so rare that we could ignore them in basically all cases anyway.

12.3 Let There be Light (in progress, optional)

Let's look a the differential form of Maxwell's equations in a space with no currents and no charge. This means we have only two nonzero equations (substituting in c, of course)

$$\vec{\nabla} \times \vec{B} = \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \tag{12.5}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{12.6}$$

Lets try to test if we can find a wavelike (ie oscillatory) solution to this. We guess that the solution looks like this

$$\vec{B} = B_0 e^{i(kz-\omega t)} \hat{\mathrm{y}}$$
$$\vec{E} = E_0 e^{i(kz-\omega t)} \hat{\mathrm{x}}$$

and propagates in the $\hat{\mathrm{z}}$ direction.

Substituting our guess into the LHS of 12.5 gives

$$\vec{\nabla} \times \vec{B} = \left(\frac{\partial}{\partial t}\hat{\mathrm{x}} + \frac{\partial}{\partial y}\hat{\mathrm{y}} + \frac{\partial}{\partial z}\hat{\mathrm{z}} \right) \times B_0 e^{i(kz-\omega t)} \hat{\mathrm{y}}$$

Notice that the only term that isn't 0 is the $\frac{\partial}{\partial z}$ term. So this simplifies to

$$\vec{\nabla} \times \vec{B} = -ikB_0 e^{i(kz-\omega t)} \hat{\mathrm{x}}$$

Now we look at the RHS of 12.5

$$\mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} = -\mu_0 \epsilon_0 i\omega E_0 e^{i(kz-\omega t)} \hat{\mathrm{x}}$$

So the whole equation gives

$$-\mu_0 \epsilon_0 i\omega E_0 e^{i(kz-\omega t)} \hat{\mathrm{x}} = -ikB_0 e^{i(kz-\omega t)} \hat{\mathrm{x}}$$

Which simplifies to

$$\boxed{\mu_0 \epsilon_0 \omega E_0 = kB_0}$$

Now we look at equation 12.6. Making our substitutions into the LHS gives

$$\vec{\nabla} \times \vec{E} = \left(\frac{\partial}{\partial t}\hat{\mathrm{x}} + \frac{\partial}{\partial y}\hat{\mathrm{y}} + \frac{\partial}{\partial z}\hat{\mathrm{z}} \right) \times E_0 e^{i(kz-\omega t)} \hat{\mathrm{x}}$$

The only nonzero term is the $\frac{\partial}{\partial z}$ term as before. Applying it gives

$$\vec{\nabla} \times \vec{E} = ikE_0 e^{i(kz-\omega t)} \hat{\mathrm{y}}$$

For the RHS, we have

$$-\frac{\partial \vec{B}}{\partial t} = i\omega B_0 e^{i(kz-\omega t)} \hat{\mathrm{y}}$$

Combining them to get the whole equation

$$i\omega B_0 e^{i(kz-\omega t)} \hat{\mathrm{y}} = ikE_0 e^{i(kz-\omega t)} \hat{\mathrm{y}}$$

Which simplifies to

$$\boxed{\omega B_0 = kE_0}$$

We can substitute the second boxed equation (solved for B_0) into the first to give

$$\mu_0 \epsilon_0 \omega E_0 = \frac{k^2}{\omega} E_0$$

Solving for $\frac{\omega}{k}$ gives

$$\frac{\omega}{k} = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

Now we just need to determine what $\frac{\omega}{k}$ means for the wave

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

where T is the period. Meanwhile k is the angular wavenumber, defined as

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

where λ is wavelength. This means that

$$\frac{\omega}{k} = \frac{\lambda}{T}$$

Since a wave travels a distance λ in each period T, this is equivalent to the wavespeed, traditionally denoted c. This means that

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$$

Plugging in the known values for ϵ_0 and μ_0, we find

$$c = 3.0 \times 10^8 \ \frac{m}{s}$$

which we recognize to be the speed of light. Now we have light from Maxwell's equations.

Part IV

Circuits

Chapter 13

Resistors and Circuit Basics

13.1 Vocabulary

There are a lot of new words we are going to need here. I'll start by defining the ones we will use immediately.

- Resistance (R): how difficult it is to make an electric current move through an object.
  - SI unit: Ohm, $\Omega = \frac{\mathrm{V}}{\mathrm{A}}$
- Conductance (G): how easily a current can pass through an object. $\mathrm{G} = \frac{1}{\mathrm{R}}$.
- Component: Any piece in an electrical circuit (other than a perfect wire).
- Resistor: A type of component that resists the flow of electric current.
- Load: Any component that consumes energy. A resistor is the simplest case.
- Voltage Source: A component (such as a battery) that creates a voltage difference between its ends even when no current flows.
- Short: Any part of a circuit where two components are connected with no resistance in between.
- Open: Any part of a circuit where no path exists between two components.
- Switch: A component that manually transitions between a short and an open.

13.2 Basics

13.2.1 Resistors and Ohm's Law

Ohm's law is an empirical relation that states that the amount of current that will flow through a circuit consisting of resistor(s) and a voltage source increases with a larger voltage difference and decreases with a larger resistance. It is only approximately true, but is good enough for our purposes.

Mathematically:

$$\boxed{I = \frac{V}{R}} \tag{13.1}$$

13.2.2 Sources

A voltage source keeps its ends at voltages that are different by whatever the voltage of the source is. For example, a 9V battery will keep its positive terminal 9V higher than its negative terminal. An ideal voltage source provides the same voltage no matter how much current would flow. Any real voltage source will decrease in voltage as its current increases due to internal resistance. In this sense, an ideal voltage source is one with 0 internal resistance.

A current source provides a fixed current through the branch. It provides the same current no matter how large the resistance becomes. Any real device will obviously not be capable of providing infinite currents. Devices are often built with fuses that are designed to fail if the current gets high enough to damage things.

13.2.3 The role of Perfect Wires

All wires in circuits are assumed to be perfect unless specifically noted. A perfect wire provides no resistance. Using $V = IR$ this means that the voltage drop across a perfect wire is always exactly 0. Put another way, **points on a circuit connected by a perfect wire will have the same voltage.** A short happens when two elements are connected by a perfect wire.

13.2.4 Short Circuits

In the case where a circuit includes a short directly between the terminals of a voltage source, we call the circuit a short circuit. In this case the total resistance is 0:

$$I = \frac{V}{0}$$

which results in an infinite amount of current. In real life, the current isn't infinite, but it can be very large. Short circuits will often destroy something in the circuit. Another way to think about this is that the voltage source is attempting to create a voltage difference between its terminals, while the wire is trying to keep those same terminals at different voltages. They can't both succeed, so something has to break.

13.2.5 Open Circuits

In an open circuit, there is no path that leads from the positive terminal of a voltage source to the negative terminal. This means that the resistance is ∞:

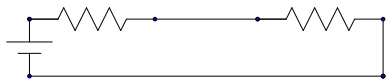
$$I = \frac{V}{\infty}$$

which results in no current. This gives a useful general result. If some branch of the circuit terminates with an open, no current will flow through that branch.

13.3 Series and Parallel Configurations

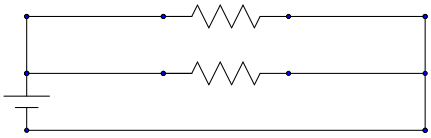
- Series: when two or more components are connected so that a current would need to travel through all of them to get from one end to the other.

  An example of a simple series circuit:



- Parallel: when 2 or more separate paths exist, so that a charge moving around the circuit must pass through exactly one of the paths.

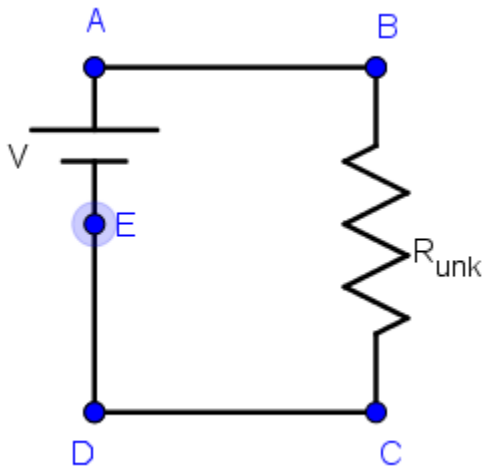
  An example of a simple parallel circuit:



13.4 Kirchhoff's Rules

13.4.1 Kirchhoff's Loop Rule

Imagine that we have the circuit shown below. We start by defining the voltage at point E, $V_E = 0$ (this immediately implies that $V_D = 0$ and $V_C = 0$ since they connect with perfect wires). Now we proceed upward from E across the battery. The battery holds the positive terminal some voltage V higher than the negative terminal. Since point A connects to the positive terminal with a perfect wire, it must also have a voltage V. Likewise, point B must have a voltage V. Going from B to C, we get some voltage drop. Recall that we already showed that $V_C = 0$. Voltage is potential energy per unit charge. The same charge at the same location better have the same potential energy, so we can't have any value except 0 for V_c. To make this happen the voltage drop across the voltage source must be V.



This result can be applied to any circuit. We call the result Kirchhoff's loop rule and state it as follows: **The sum of the voltage changes across a closed loop is exactly 0**. Mathematically

$$\boxed{\sum \Delta V = 0}$$

Remember that this only applied when going all the way around a closed loop. An equivalent statement is: potential energy of charges in a circuit is well defined.

13.4.2 Kirchhoff's Junction Rule

We know that charge must be conserved. Referring again to our circuit above, any charge that flows into point A must either stay at point A, or move through the circuit. If a bunch of charge moved to A and

stayed there, the result would be a huge repulsive force that would push the charge away. This means that all the charge that flows into A has to flow out. This gives us Kirchhoff's Junction Rule. **The sum of the currents into and out of a node is exactly 0.** Mathematically

$$\boxed{\sum I = 0}$$

an equivalent statement is: charge in a circuit is conserved

13.5 Adding Resistors

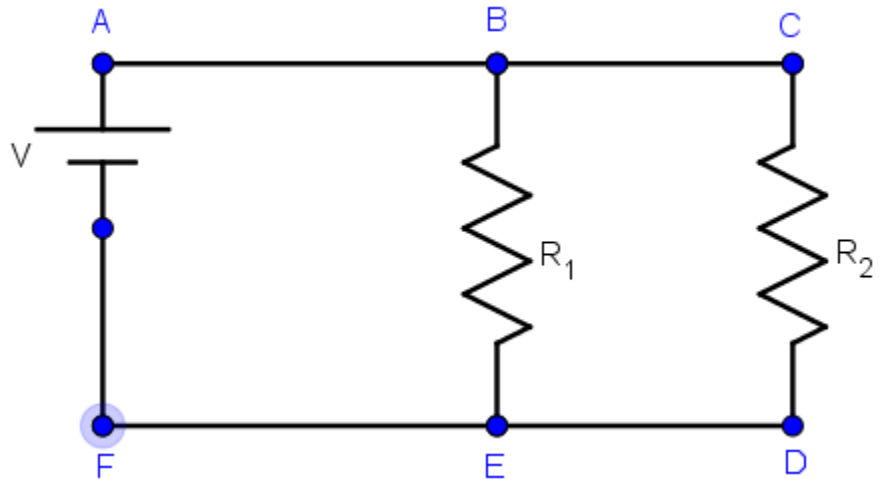
13.5.1 Series

Imagine that I have a resistor with a resistance of R. I cut it in half and then put both halves of the resistor back into the circuit with only a perfect wire between them. Clearly the resistance of the circuit should not have changed. This tells us that $\frac{R}{2} + \frac{R}{2} = R$. I doubt that this is particularly surprising. We could cut R into as many pieces as we want to get whatever ratios we desire. This means that for resistors in series, we can simplify the circuit by adding all resistors that are in series together to get a single large resistor.

$$\boxed{R_{eq} = R_1 + R_2 + ...}$$

13.5.2 Parallel Resistors

Consider the parallel circuit shown



In this circuit, R_1 and R_2 are in parallel with each other because charge cannot move around unless it passes through either R_1 or R_2, and no charge will pass through R_1 and R_2.

Because A,B,C are connected by perfect wires to the source's positive terminal, they have voltage V. Since F,E,D are connected by perfect wires to the source's negative terminal, they have voltage 0. This means that R_1 and R_2 both have a voltage V across them from the source.

This means that the current that will flow through R_1 is

$$I_1 = \frac{V}{R_1}$$

and through R_2 we have

$$I_2 = \frac{V}{R_2}$$

Then the total current that flows through the circuit is

$$I_{tot} = I_1 + I_2 = \frac{V}{R_1} + \frac{V}{R_2}$$

We want to find a single resistor that would be equivalent to the two resistors in parallel. For this equivalent resistor, we would have

$$I_{eq} = \frac{V}{R_{eq}}$$

It wouldn't be equivalent unless $I_{eq} = I_{tot}$. So we have that

$$\frac{V}{R_{eq}} = \frac{V}{R_1} + \frac{V}{R_2}$$

Which gives the general law

$$\boxed{\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + ...}$$

13.6 Solving Resistor Circuits

You can either go directly to Kirchhoff's laws, or do some simplification first. The simplification way looks something like the "procedure" below. At any point if you don't see how to continue, you can always just write out Kirchhoff's laws at that points.

- If the circuit is $3 - D$ or laid out in a bizarre way, make it nice.
- Remove any resistors that have ends with the same potential. No current may flow through them.
- If the circuit has any resistors in series or parallel, replace them with equivalent resistors
- Find the current through the whole circuit by using the equivalent resistor circuit.
- Now that you know the current, return to a less simplified circuit to find whatever you were supposed to find.

For some complicated circuits, you may need to repeat previous steps if other simplifications made it clear that you missed something. For example, after some simplification you might notice that two points in a circuit have the same voltage, making resistors between those points irrelevant.

13.7 Power Dissipated in Circuits

We know that $P = \frac{dW}{dt}$ for a circuit we could write

$$\frac{dW}{dt} = \frac{dW}{dq}\frac{dq}{dt}$$

or

$$\frac{dW}{dt} = VI$$

which is usually written as

$$\boxed{P = IV} \tag{13.2}$$

For a resistor, $V = IR$, so this can be written as

$$P = IV = I^2R = \frac{V^2}{R}$$

All of which are useful depending on what you are given. The boxed form is the definition, and the others are derived from Ohm's law. This means only the boxed form works for non-resistor elements.

Chapter 14

Reactive Circuits

We end the course with a discussion of reactive circuit elements: inductors and capacitors.

14.1 Capacitors

Imagine that we are looking at a system consisting of two parallel plates next to each other and hold the two plates at different voltages. An electric field must exist between the plates. This implies that positive charge has collected on the plate connected to the higher voltage, and negative charge has collected on the plate connected to the lower voltage.
If we made the plates much larger in area, we would have the same difference in potential, and the same field between them, but more charge on each plate.

If the voltage difference is V, it's not hard to find how much charge will accumulate on the plates. Take a Gaussian surface consisting of a cylinder with one face between the plates and the other face outside the higher voltage plate. Ignore the lower voltage plate for the moment, we will use superposition to get it later.

The plates are much wider than their separation, so we will pretend they are infinite. We've done this problem a few times before. The field is $\frac{\sigma}{2\epsilon_0}$ and points toward the low voltage plate. The other plate will have the same field, and will also point toward the low voltage plate. So the total field between the plates is

$$E = \frac{\sigma}{\epsilon_0}$$

outside the plates, the fields are in opposite directions and cancel to 0. The total charge can be written as σA where A is the plate areas. Since the field is constant, we can say that $E = \frac{V}{\Delta x}$ where Δx is the separation distance

$$EA = \frac{Q}{\epsilon_0}$$

$$\frac{AV}{\Delta x} = \frac{Q}{\epsilon_0}$$

so the charge Q will be

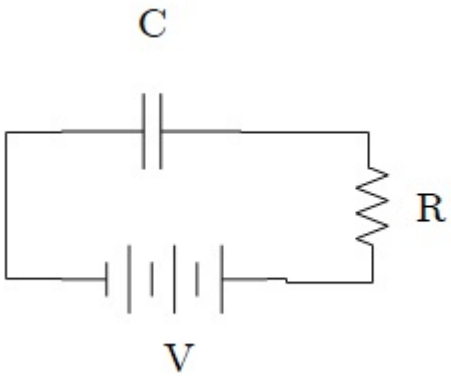
$$Q = \frac{A\epsilon_0}{\Delta x} V$$

We can always take a combination of constants and combine them into one for simplicity. It is conventional to do that here.

$$\boxed{Q = CV} \tag{14.1}$$

Note that C is actually defined to be the ratio $\frac{Q}{V}$, so this definition actually works for any geometry. Its better to think about C as a function only of the geometry though. In particular, cylindrical capacitors are very common in practice.

Now that we have the idea of a capacitor, we can think about what the charging might look like. Consider a circuit like this



Assume that the source V is shut off initially and is turned on at time $t = 0$. We can use Kirchhoff's loop law to say

$$V = V_c + V_r$$

or

$$V = \frac{Q}{C} + IR$$

We know that $I = \dot{Q}$

$$V = \frac{Q}{C} + \dot{Q}$$

Guess that $Q = VC + Q_0 e^{kt}$ if you hadn't included the VC, you would have just had more algebra later, but it would still work
plugging that in gives

$$V = V + \frac{Q_0}{C} e^{kt} + Q_0 kRe^{kt}$$

Which simplifies to

$$-\frac{1}{RC} = k$$

So our solution is

$$Q = VC + Q_0 \exp\left(\frac{-t}{RC}\right)$$

Obviously, the capacitor charge must be 0 at $t = 0$, since it hasn't started charging yet so $Q_0 = -VC$ and our whole expression becomes

$$Q = VC\left(1 - \exp\left(\frac{-t}{RC}\right)\right)$$

Which you can easily check goes to VC at long time and is 0 at $t = 0$ as needed.
So capacitors charge exponentially. You might guess that by symmetry, they should discharge exponentially.

If we take a derivative of this to get the current, we get

$$I = \frac{V}{R} \exp\left(\frac{-t}{RC}\right)$$

Note that this goes to 0 at long times once the capacitor has charged. Intuitively, no more current can flow through a circuit once the capacitor charge reaches equilibrium. This implies that for a pure DC voltage,

a capacitor is an open circuit. A related note is that before the capacitor has charged at all, there is no voltage across it. This means that an uncharged perfect capacitor acts as a short with respect to sudden voltage changes.

All this gives us a key piece of intuition. Capacitors function as integrators of the current for time. This results directly from the simple observation that $Q = \int I \, dt$. Consider a circuit with a sinusoidal current. Then the voltage across the capacitor would be the integral of a sine, which is a cosine. Cosine is always $90°$ out of phase with sine. This means that the current through the capacitor is $90°$ out of phase with the voltage across the capacitor. This turns out to be a general property of capacitors and has some interesting implications for AC circuits.

14.1.1 Energy Storage

A capacitor can be thought of as a storage device for electrostatic energy. Recall that we can get the total energy of a system using

$$U = \frac{\epsilon_0}{2} \int E^2 \, d\tau$$

where the integral is over all space. This is considerably easier for a capacitor because the field is only non-zero between the plates, and it is constant there The field will be

$$E\Delta x = \frac{Q}{C}$$

so

$$E^2 = \frac{Q^2}{C^2 \Delta x}$$

so our energy equation is (noting that the volume integral is just multiplication by volume)

$$U = \epsilon_0 \frac{Q^2}{2C^2 \Delta x^2} A\Delta x$$

Use that $V = \frac{Q}{C}$

$$U = \frac{\epsilon_0}{2\Delta x} V^2 A$$

But C for a parallel plate capacitor (which is what we are using for this derivation) is $C = \frac{A\epsilon_0}{\Delta x}$, so this simplifies to the very nice expression

$$\boxed{U = \frac{1}{2} C V^2} \tag{14.2}$$

Note that regardless of what geometry we choose we always get the same expression. The constant C is just different for different geometries.

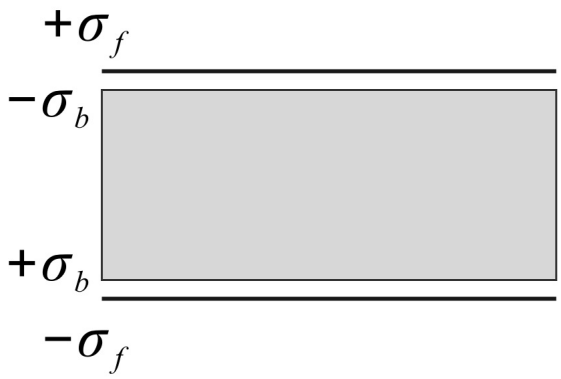
14.1.2 Dielectrics in Capacitors

A dielectric is any substance that is an insulator and can be polarized by an electric field. We will only treat linear isotropic dielectrics here. Imagine that we put something polarizable in a field. We would expect charge to appear on each surface of the object. The amount of charge would depend on the field there, which would be the sum of the input field and the induced field. What is really happening, is that every point is polarizing, but that the charges cancel everywhere except the ends. This is fairly abstract, but will be more concrete if we consider a capacitor.

Imagine that a dielectric is placed inside a charged capacitor. As a result of the field, the dielectric will polarize. Because the charges cannot move freely, they will just shift over slightly. In the middle of the

Figure 14.1: A

dielectric inside a charged capacitor. There will be a net charge on the top and bottom of the dielectric. Because negative charges are attracted to positive and vis-versa, the direction of the field created by the bound charges is opposite to that created by the charge on the capacitor plates.



dielectric, this all cancels, but at the end, charge is left on both ends as shown in diagram 14.1. This will result in the creation of a second field from the bound charges that opposes the original field.
To determine the field, we start with the reasonable assumption that the amount of bound charge will be proportional to the total field, which will be the sum of the created field and the original field.

$$E_{tot} = E_f + E_b$$

where

$$E_f = \frac{\sigma_f}{\epsilon_0}$$

$$E_b = \frac{\sigma_b}{\epsilon_0}$$

so

$$E_{tot} = \frac{\sigma_f - \sigma_b}{\epsilon_0}$$

where the negative sign is because the fields are in opposite directions.
We now use the assumption that the bound charge depends on the total field. We guess that the constant of propotionality is ϵ_0 (to get the units right) times some dimensionless constant χ

$$\sigma_b = \epsilon_0 \chi E_{tot}$$

Then

$$E_{tot} = \frac{\sigma_f - \epsilon_0 \chi E_{tot}}{\epsilon_0}$$

Solving gives

$$E_{tot} = \frac{\sigma_f}{\epsilon_0(1+\chi)}$$

Now imagine that the plates of the capacitor are isolated as we insert the dielectric. This prevents any charge from leaving so that the Q on each capacitor remains the same.
Recall that

$$Q = CV$$

This means, for capacitor before insertion,

$$Q = C_0 V_0$$

after insertion

$$Q = C_d V_d$$

Now we divide the capacitor equation after, by the capacitor equation before to get

$$1 = \frac{C_d}{C_0}\frac{V_d}{V_0} \tag{14.3}$$

But, since

$$V = \int \vec{E} \cdot d\vec{l}$$

Since the field is constant, this is just

$$V = Ed$$

the distance is unchanged, so

$$\frac{V_d}{V_0} = \frac{E_d}{E_0} = \frac{1}{1+\chi}$$

so, returning to 14.3

$$1 = \frac{C_d}{C_0}\frac{1}{1+\chi}$$

which means that our capacitor now has a capacitance of

$$C_d = C_0(1+\chi)$$

This is usually written in terms of ϵ_r instead. The relation between χ and ϵ_r is

$$\chi = \epsilon_r - 1$$

so we could write that

$$\boxed{C_d = C_0\epsilon_r} \tag{14.4}$$

For a parallel plate capacitor, we know

$$C_0 = \epsilon_0 \frac{A}{d}$$

For our dielectric capacitor, this would be

$$\boxed{C_d = \epsilon_r\epsilon_0\frac{A}{d} = \epsilon\frac{A}{d}} \tag{14.5}$$

For the second equality, the definition $\epsilon = \epsilon_r\epsilon_0$ was used. Equation 14.4 gives the dielectric filled capacitance in terms of the vacuum filled, but we still could have used only ϵ like we did in equation 14.5, but we would need to know the capacitance for that particular geometry.

14.2 Inductors

Recall that mutual inductance (see section 11.1) referred to the tendency of changing current in one object to induce voltage in another. It is also possible for changing current in an object to create an induced voltage in itself. This is called self-inductance and is denoted L. Like the mutual inductance, self inductance is purely geometric. An easy way to get very high self inductance is to have many loops in a row. We will be more interested in how the object behaves in a circuit. Recalling from our inductance discussions earlier, we expect there to be an induced voltage in the inductor that opposes the change in flux. In circuits, this flux is a result of the magnetic field the inductor produces as a current is passed through it. This basically means that inductors will resist changes in current. If we define self inductance like mutual inductance, we should have

$$\boxed{\mathcal{E} = -L\frac{dI}{dt}} \tag{14.6}$$

Since nothing in the theory of induction has changed, we will focus on how they work in circuits.

14.2.1 Inductors in Circuits

If an inductor resists changes in current, we expect that very high frequency signals will be completely blocked, while DC signals will be unaffected. Note that this is exactly the opposite behavior that the capacitor had. The combination of the two elements are the basis for analog signal filtering.

Lets imagine that we have some circuit (the details aren't important), that has a current of

$$I = I_0 \cos \omega t$$

. By using equation (14.6), we can see that the voltage across the inductor is

$$\mathcal{E} = LI_0\omega \sin \omega t$$

We know that power is $P = IV$. Here I and V are functions of time, so we will have to integrate to get the energy dissipated. Technically the V here is the voltage drop across the element, so that is $V = -\mathcal{E}$. Still, this won't be very important. Let's start by integrating over one complete period. If the angular frequency is ω, then the period should be $T = \frac{2\pi}{\omega}$. I will just use T in the bound for simplicity. This integral is

$$E = -LI_0^2\omega \int_0^T (\cos \omega t)(\sin \omega t)dt$$

This gives 0. The element doesn't dissipate power. But now lets try another integral.

$$E = -LI_0^2\omega \int_0^{\frac{T}{4}} (\cos \omega t)(\sin \omega t)dt$$

This gives

$$E = -\frac{1}{2}LI_0^2$$

This means that the element gave the circuit energy... But where did it come from? To see that, let's evaluate another integral

$$E = -LI_0^2\omega \int_{\frac{T}{4}}^{\frac{T}{2}} (\cos \omega t)(\sin \omega t)dt$$

This gives

$$E = \frac{1}{2}LI_0^2$$

Now we can understand that the inductor was just giving the circuit back the energy that it took earlier!

What was going on in the inductor at that time? The magnetic field in the inductor should be proportional to the current (as with any wire assembly). This means that the magnetic field in the inductor should peak when the current peaks. In our case that was at $t = 0$. Then it will reach zero magnetic field at $t = \frac{\pi}{\omega} = \frac{T}{4}$, which is exactly when the inductor finished giving back the energy it had.

We already knew that electromagnetic fields store energy, so maybe it wasn't too surprising that a device could exist that could temporarily store magnetic energy. Either way, we now understand the function of an inductor. Inductors store energy (ultimately in the magnetic field) and then give it back later. If both capacitors and inductors function as energy storage devices, what happens when we combine them? Think about that for a minute before proceeding to the next section, where we will do just that.

14.3 Oscillatory Circuits

Lets say that we make a circuit that consists only of a perfect inductor, a perfect capacitor, a switch, and perfect wires. Then Kirchhoff's loop rule says

$$L\frac{dI}{dt} + \frac{Q}{C} = 0$$

It will be easier to put everything in terms of Q

$$L\ddot{Q} + \frac{1}{C}Q = 0$$

or

$$\ddot{Q} = -\frac{1}{LC}Q$$

Wait a minute... We've seen this equation before... It's the spring equation... We guess that

$$Q = Q_0 \exp(i\omega t)$$

plugging this guess in gives

$$i^2\omega^2 Q = -\frac{1}{LC}$$

$$\omega = \sqrt{\frac{1}{LC}}$$

We could always start our time when the capacitor was fully charged, so we keep only the cosine.

$$Q = Q_0 \cos\sqrt{\frac{t}{LC}}$$

This is remarkable. Our circuit behaves like a perfect spring. We also notice something. If Q is oscillatory, then so is the energy in the capacitor $U = \frac{1}{2}CV^2$ which we can rewrite using $Q = CV$ to $U = \frac{Q^2}{2C}$.

Since it need to return to it's original value, it can't be lost. I must go into the inductor. It does. It was stored in the electric field in the capacitor, but in the magnetic field in the inductor. We can verify explicitly that energy is conserved by using that $U = \frac{1}{2}LI^2$ for inductors.

Finding the current explicitly gives

$$I = -\frac{Q_0}{\sqrt{LC}}\sin\frac{t}{LC}$$

I'll keep the argument of sine and cosine in the form ωt for clarity Plugging in Q for the capacitor gives

$$U_c = \frac{Q_0^2}{2C}\cos^2\omega t$$

Plugging in I to the inductor energy equation

$$U_L = \frac{1}{2}L\frac{Q_0^2}{LC}\sin^2\omega t = \frac{Q_0^2}{2C}\sin^2\omega$$

$$U_t = U_L + U_c = \frac{Q_0^2}{2C}(\sin^2\omega t + \cos^2\omega t)$$

or

$$U_t = \frac{Q_0^2}{2C}$$

So we have a remarkable fact. LC circuits are electromagnetic oscillators. Instead of transferring energy back and forth between spring potential and kinetic, this oscillator transfers it between electric fields in the capacitor and magnetic fields in the inductor.

Now try to think about what should happen if we add a resistor. You solved that case for a spring-mass. Does it carry over?

Chapter 15

Closing Thoughts

There is obviously still a lot left to be done here. I need to provide example problems for every topic listed here. I need to add sections for some minor topics. In some areas, the level of the math could be brought up or down for consistency. Still though, I hope you took away the following lessons

1. Math equations are a way of expressing concepts. Trickier concepts require trickier math.
2. The same math ends up showing up in very different areas, but still refers to the same concepts.
3. Physics is concerned with finding the smallest set of equations that explain the widest possible set of scenarios.