lecture 1

Topics:

What is physics?
What is classical mechanics?
Degrees of freedom
The Art of Theoretical Physics
Motion, trajectories and $F = m\,a$
$F = ma \Leftrightarrow$ two initial conditions/degree of freedom
Finding trajectories numerically
Forces of the form $F(t)$

What is physics?

According to Wikipedia

> Physics (from the Greek, $\phi\upsilon\sigma\iota\zeta$ (physis), "nature" and $\phi\upsilon\sigma\iota\kappa o\zeta$ (physikos), "natural") is the science of Nature, from the quarks to the cosmos. Consequently, physics treats of the fundamental constituents of the universe, the forces they exert on one another, and the results produced by these forces. Sometimes, in modern physics, a more sophisticated approach is taken that incorporates elements of the three areas listed above; it relates to the laws of symmetry and conservation, such as those pertaining to energy, momentum, charge, and parity.

As much as I like symmetry and conservation laws (which we will certainly talk a lot about in this course) I am not sure that I like this definition much. It focuses a little too much on grand theoretical principles and not enough on how useful physics is. Both poles of physics - the theoretical and the experimental/applied are crucial to the way physics works. True - physics addresses the most fundamental questions in the universe - but we can address these questions usefully only because we can use physics to build the tools required to answer the most difficult questions. And the answers to these questions, in turn, allow us to build better tools and ask deeper questions. This self-containedness - this cyclic give and take between mathematical theory and tool-building is the glory of physics itself, and the reason why physics is important to all other areas of science. Physicists build tools, both technological and mathematical, that are crucial to the pursuit of scientific questions in all fields and which form the basis for much of our mathematics and technology.

And besides, physics is just a marvelously fun way of looking at the world.

What is classical mechanics?

Classical mechanics is a category defined by what it is not! It is not quantum mechanics. A **classical mechanical system** is any collection of objects that we can describe to a good approximation without worrying about quantum mechanics. This includes most of the systems you are used to in everyday life (if you don't worry too much about what goes on inside your computer or CD player or TV set). In this course, we will discuss these mechanical systems, and we will push beyond your everyday experience to discuss what happens when objects move at speeds close to the speed of light. We won't study quantum mechanics explicitly, but we will talk about it at times. The underlying laws of the Universe seem to be quantum mechanical, and it is fun, instructive, and not at all trivial to think about how classical physics emerges as an approximation to the quantum mechanical world.

Degrees of freedom

We will start slowly, with some useful definitions.

Coordinates — The coordinates of a physical system are the numbers (possibly dimensional) that describe the system at a given fixed time.

Frame of reference — The coordinates we use to describe the system are always somewhat arbitrary, depending on conventions. Mathematically, we should be able to change our system of coordinates and still describe things. But in physical problems, it is usually easiest to restrict the allowed coordinate systems. Fortunately, in physics, there is an obvious way of specifying the conventions that we use in a physical problem. We simply describe in detail how to measure the coordinates at any given time. Such a description is called a **frame of reference.** We will not always mention the **frame of reference** explicitly, because we have in mind something very intuitive and simple - the set of clocks, measuring sticks and other stuff that we need to study classical phenomena in a laboratory. But we will see (particularly when we discuss relativity in a few weeks) that it is important to **have** a frame of reference that fixes our conventions. Otherwise, we can get into all kinds of problems.

Configuration — The **configuration** $q(t)$ of a mechanical system is a number or vector consisting of values of a set of independent **coordinates** that completely describe the system at time t.

We will call the configuration q, without specifying (at least for another few seconds) whether q is a single variable, or some kind of vector describing several coordinates at once. The word "independent" in the definition means that none of the numbers in our set of coordinates is redundent or dependent of the others. That is we asume that the coordinates in the configuration are independent in the sense that each can in independently changed, and each different set of values describes something physically different. Given a set of values of the coordinates at some particular time, we can figure out what the configuration is and thus **what the system looks like** at that time. Thus a configuration is just a mathematical snapshot of the system at a given time.

One way of describing the underlying problem of classical mechanics is that we want to understand how the configuration of the system evolves with time. That is we want to put the snapshots together into a mathematical movie to describe how the system moves.

Degrees of freedom — The number of independent components of the configuration q is called the number of **degrees of freedom** of the mechanical system.

The number of degrees of freedom is the number of independent ways in which the system can move. Here are some examples. A point mass sliding on an airtrack, described by only a single coordinate, has one degree of freedom. We get more degrees of freedom if we go to more dimensions or to more complicated objects.

system	coordinate	# of DOFs
point mass on a track	ℓ (distance along the track)	1
point mass on a flat surface	(x, y)	2
point mass in 3-d	$\vec{r} = (x, y, z)$	3
rigid body in 3-d	$\vec{r}$ of center + 3 angles	6
2 masses + massless spring	$\vec{r}_1$ and $\vec{r}_2$	6
2 masses + massive spring	$\vec{r}_1$, $\vec{r}_2$ and spring	"∞"

A continuous massive spring formally has an infinite number of degrees of freedom, because to specify its configuration we would have to give a continuous function describing how much every point on the spring is stretched. Really, of course, a physical spring has a finite but very large number of degrees of freedom, because it is not actually continuous, but is made up of atoms. But the difference between ∞ and Avogadro's number is often not very important.

This brings up an important philosophical point. What the heck is a "point" mass? What is a "rigid" body? What is a "massless" spring? Most of you have probably been dealing with physics problems for so long that you are used to these phrases. But it is important to remember that these are mathematical idealizations. Real physical systems are complicated, and in fact, what we choose for q may depend on what kind of physical questions we want to ask, what level of accuracy we need in the answer, and even how long we want to study the system.

So for example, for a hockey puck sliding on the ice at the Boston Garden, we might decide that the configuration is specified by giving the x and y coordinates that determine the puck's position in the plane of the ice. Then q would stand for the two dimensional vector, (x, y). But if we do this, we have ignored many details. For example, for a shot that comes off the ice, we would need to include the z coordinate to describe the motion of the puck. For some purposes, we would also need to include descriptions of the puck itself. For example, we have not included an angular variable that would allow us to specify how the puck is turned about its vertical axis. This is probably good enough for most problems. But sometimes, more information is required to give a good description of the physics. For example, if we wanted to understand how a rapidly rotating puck moves, we might need this more detailed information. We could also go on and describe how the puck might deform when hit by the stick, and so on. We could include more and more information until we got down to the level where we begin to see the molecular structure of the rubber of the puck. At this point, we begin to see quantum mechanical effects, and classical mechanics is no longer enough to give an accurate description.

The Art of Theoretical Physics

This is a good lesson. The coordinates that we use to describe the system may depend on what kind of information we want to get out of our mechanical model of the system, and how accurately we want our model to reproduce reality. We usually will not go over these niceties each time we discuss a system, but they are important to remember. There is really a very important point here. In physics courses, we frequently discuss "toy" systems which are obviously oversimplified, in which we have clearly left out features that are important in the "real world." This is not something to apologize for. This is precisely the art of theoretical physics. We work hard to abstract the essential physics of a system, without including things that don't matter at the level of description that we are interested in. This down-to-earth ability to focus on the crucial parameters is far more important than fancy mathematical gymnastics.

In fact, I believe, getting better at this art is one of the most useful things you can get out of this course. It is generally useful far beyond this or future physics courses. The ability to build mathematical models of phenomena is crucial to many fields. But models can be as misleading as they can be useful unless they focus on the right parameters, and unless the modeler is aware of the model's limitations. Physics is the paradigm for this kind of thinking. This is one of the reasons why, over the years, trained physicists have been so much in demand in very different fields.

Motion, trajectories and $F = m\,a$

Trajectory — A **trajectory** is a possible **motion** of a physical system.

We describe a trajectory by giving the value, $q(t)$ of the coordinates of the system as a function of time. There are an infinite number of possible trajectories for any interesting physical system (that is, unless everything is nailed down and completely fixed). Our everyday experience tells us that knowing the configuration of the system at some time is not enough information to allow us to calculate the trajectory. A snapshot doesn't tell us how the system is moving. We also need to know the velocity of the system — the derivative of q with respect to time. We will sometimes use the notation of a dot over an object to represent a time derivative:

$$\dot{q}(t) \equiv \frac{d}{dt} q(t) \,. \tag{1}$$

If q describes more than one coordinate, then $\dot{q}$ describes the same number of independent velocities - but we will continue to call $\dot{q}$ the velocity, even if it has multiple components. If q is a vector, $\dot{q}$ is the same kind of vector. For the hockey puck,

$$q = (x, y)\,, \qquad \dot{q} = (\dot{x}, \dot{y})\,. \tag{2}$$

Then we can say that if we know q and $\dot{q}$ at a particular time, we should be able to calculate the trajectory. For example, if we know where the hockey puck is and how fast and in what direction it is moving, we can plot its trajectory by just assuming that it continues to move at a constant speed along the line determined by its position and velocity.

Something important has happened. I have snuck in one of the critical assumptions of mechanics. Let me say it with the formality it deserves.

> $\mathcal{A}$ **— The trajectory, or subsequent motion, of a mechanical system is completely determined if we know the configuration,** q **and the velocity,** $\dot{q}$ **at some given time.** (3)

Because the number of components of q and of $\dot{q}$ are both equal to the number of degrees of freedom, this means that we have to specify **two constants per degree of freedom** in order to specify how the system moves. This is a biggie. We will come back to this again and again, as you will see.

One of the central problems of mechanics is to calculate what these trajectories are for various systems — that is to say, to figure out how things move. To begin, let us think about what this process is like for a system specified by a single coordinate, x. Newton tells us that to figure out how x changes as a function of time, we define our coordinates in an appropriate inertial frame (we'll talk much more about what this means later) and use his second law

$$F = m\,a \tag{4}$$

where F is the force on the object, m is its mass, and a is its acceleration (also sometimes written as $\ddot{x}$ — one dot for each time derivative)

$$a \equiv \frac{d^2}{dt^2} x \equiv \ddot{x} \tag{5}$$

We will have much more to say about mass, but for now, we will just assume that m is a fixed property of the object. The force, F, is more complicated. In general, F will depend on what the system is doing. Since we have already assumed that all we need to know about the system at a given time is q and $\dot{q}$, in this case x and $\dot{x}$, all F can depend on is q, $\dot{q}$ and t. So in general, F at time t is some function of x and $\dot{x}$, at that same time, and t, $F\Big(x(t), \dot{x}(t), t\Big)$. Then Newton's second law becomes a formula for the acceleration:

$$\ddot{x} = \frac{1}{m}\, F(x, \dot{x}, t) \tag{6}$$

Note that in general, in Newton's second law, the force is a function of all three variables, x, $\dot{x}$, and t. Because of things like this, we will often have to discuss the calculus of functions of several variables. I realize that some of you are just beginning to study this subject formally in math courses. In general, in this course, I will often make use of mathematics that many of you have not seen. We won't be using any deep properties of the subject, just things that you would guess immediately from your knowledge of calculus of a single variable, vectors, and algebra. The proofs can wait until (or if) you take the appropriate math course. The important thing is that you try to understand the physics, and not faint when unfamiliar mathematics is thrown at you.

$F = ma \Leftrightarrow$ **two initial conditions/degree of freedom**

Now the first deep question I want to address is this: Why should the fundamental law of mechanics (which Newton's second law certainly is) be a formula for acceleration? A partial answer comes from the mathematics. A formula like (6) is called a **second order differential equation** because it involves a second derivative but no higher derivatives. Our mathematician friends tell us that a second order differential equation for x has in general an infinite number of solutions, which can be labeled by the values of x and $\dot{x}$ at some given time. These conditions that specify the solution are sometimes called the **initial conditions**, and this is certainly appropriate for the physics applications we have in mind. Once we specify the initial conditions, the initial position and velocity, the solution of the differential equation, which is the trajectory, is completely specified. This means that if the fundamental law is a second order differential equation, like (6), the mathematics guarantees that assumption $\mathcal{A}$ is satisfied. Thus we have shown that $F = ma$ implies $\mathcal{A}$. This is a good thing.

In general, I am not so interested in proofs in this course, at least not in the mathematical sense. However, in the notes, and in lecture if I have time, I will give a "physicst's proof" that the solution of a second order differential equation is specified by q and $\dot{q}$.

Here is a simple and probably very familiar example. Suppose that the force is constant,

$$F(x,\dot{x},t) = F_0 = m\,a = m\,\frac{d^2}{dt^2}x = m\,\ddot{x} \tag{7}$$

The solution for $x(t)$ can be written in terms of x and $\dot{x}$ at $t = 0$ as

$$x(t) = \frac{F_0}{2m}t^2 + \dot{x}(0)\,t + x(0) \tag{8}$$

We will see one general way of getting this result later. But is easy to check that (6) is satisfied by computing the second derivative.

$$\dot{x}(t) = \frac{F_0}{m}t + \dot{x}(0) \tag{9}$$

$$\ddot{x}(t) = \frac{F_0}{m} \tag{10}$$

This shows that the t^2 term is right. You can check that the other terms in (8) are correct by setting $t = 0$ in (8) and (9).

The latest version of Mathematica has a fantastic function that allows you to input initial conditions in real time. Check out the notebook **lecture-1-1.nb**.

There is nothing special about $t = 0$. We could just as well write the solution in terms of the position and velocity at some arbitrary time t_0, as

$$x(t) = \frac{F_0}{2m}(t-t_0)^2 + \dot{x}(t_0)\,(t-t_0) + x(t_0) \tag{11}$$

I hope that I have convinced you that $F = ma$ implies two initial conditions/degree of freedom. That doesn't mean that Newton's second law is equivalent to $\mathcal{A}$ because we haven't show that any system that requires two initial conditions/degree of freedom satisfies $F = ma$. But there is a mathematical sense it which this true at least for one degree of freedom. I am not going to talk about this in lecture, because the argument is a slightly intricate exercise in multivariable calculus. But it may be fun for those of you who like that sort of thing, so I have included it as an appendix to the lecture notes.

What have we really learned from this mathematical philosophy? I have argued that Newton's second law is actually equivalent to assumption $\mathcal{A}$ — each implies the other. Of course, we haven't really understood why F is equal to ma, because we still don't know why assumption $\mathcal{A}$ is true.

That leads naturally to another question. Must $\mathcal{A}$ be true for some reason? Can we imagine a world very different from our own in which it is not? Actually, I don't think $\mathcal{A}$ must be true. In fact, I don't really think it is true. Like most of the things we teach as "physical laws", it is only an approximation to what is really going on. We will come back to this much later. But this is a nice example of the kind of thinking that we will be doing a lot of in this course.

Should you take Physics 16?

Because this is the first day of the class, you are all full of adrenaline and are paying close attention to every word I say. But normally, at about this point in a 1.5 hour lecture, you would all be falling asleep and would need a little something to perk you up. This is where we normally put a **miniexam**, a small problem that you work on with a random group of other students. Since we haven't done enough yet to warrant a miniexam, I thought that instead, I would say a bit about the course, and then see whether you have any questions and maybe ask you some questions myself. There is a lot more detailed information on the information sheet on the web page. I work hard to keep it updated, so please read it carefully.

Presumably you are all here because you have good preparation in physics and math and you are afraid that Physics 15a would be boring. The first thing to say is that Physics 15a would probably not be boring. It is a good rigorous course, and almost all of you would learn something. However, for many of you, Physics 16 may be more fun.

As you may have guessed from what has happened so far, I will spend a lot less time on statics, assuming you already know about adding up vector forces and torques on objects that are not moving. We have already started looking at things moving around. In equations, I won't spend time thinking about $\vec{F} = 0$. We are going right to $\vec{F} = m\,\vec{a}$.

This will give us some time to do cool things that are not done in a good conventional college physics course like Physics 15a, including the following:

1. linearity and normal modes
2. Lagrangian mechanics and Noether's theorem
3. relativistic strings
4. the moment of inertia tensor
5. inflation before the big bang, dark matter and dark energy

We won't explore any of these ideas in the depth that they deserve. For that you will have to take more advanced physics courses. But neither will we look at them qualitatively or superficially. From each of these subjects, we will try to extract some beautiful and important pieces that we can analyze honestly and in detail.

I would make one request. Don't take 16 because you think that it is going to be hard for you and you want a challenge. And don't drop it because you think it is going to be hard for you and you don't want a challenge. In fact, Physics 16 is in many ways a very humane course. We will do civilized things like having a take-home component of our exams so that we can correct for problems that some people have with time pressure on exams. And while the material is hard, as long as you get the help that we are eager to give you, the work is really not very different from that in 15a. For some of you it may take a little more time, so as I say at the end of the information sheet, it is important to think about it in the context of your whole schedule, curricular and extracurricular.

Finding trajectories numerically

If you are given a complicated force law, depending (for one degree of freedom) on x, $\dot{x}$ and t, you will usually not be able to solve for the trajectories in terms of known functions. However, you can always find them by brute force, solving $F = ma$ numerically. One simple way to do this is to make small time steps and use the most important formula in Physics — the Taylor expansion. Suppose that you are given the initial conditions, $x(t_0)$ and $\dot{x}(t_0)$. Then you can find x and $\dot{x}$ approximately for a slightly later time, $t_0 + \Delta t$, by Taylor expanding the function $x(t_0 + \Delta t)$ around t_0. The Taylor expansion looks like this:

$$x(t_0 + \Delta t) = x(t_0) + \Delta t\, \dot{x}(t_0) + \cdots \tag{12}$$

where the $\cdots$ are terms with more powers of Δt. I love the Taylor expansion, and we are going to come back to it many times in the course, and later we will talk about the higher order term. But here, these can be ignored if Δt is sufficiently small. This first term in the Taylor expansion is nothing more than the definition of the derivative:

$$x(t_0 + \Delta t) = x(t_0) + \Delta t\, \dot{x}(t_0) + \cdots \quad \Leftrightarrow \quad \dot{x}(t_0) = \frac{x(t_0 + \Delta t) - x(t_0)}{\Delta t} + \cdots \tag{13}$$

How does this help us find $x(t)$. This is fine as far as it goes, because we are given the initial conditions for $x(t_0)$ and $\dot{x}(t_0)$. But what we would like to do is to be able to iterate this procedure and find $x(t_0 + n\Delta t)$ which would mean that we could find $x(t)$ (at least approximately) at a whole sequence of future times. And in the next step we have a problem because we don't know $\dot{x}(t_0 + \Delta t)$ from the initial conditions. But now the key point is that we can also keep track of $\dot{x}$

$$\dot{x}(t_0 + \Delta t) = \dot{x}(t_0) + \Delta t\, \ddot{x}(t_0) + \cdots \tag{14}$$

And we know $\ddot{x}(t_0)$ from $F = ma$:

$$\ddot{x}(t_0) = a(t_0) = \frac{1}{m} F(x(t_0), \dot{x}(t_0), t_0) \tag{15}$$

You can always do this for $F = m\,a$. Given x and $\dot{x}$ at some time t_0, you can always find x and $\dot{x}$ for an infinitesimally later time, $t_0 + \Delta t$.

Now that you know x and $\dot{x}$ for the later time $t_0 + \Delta t$, you can apply exactly the same procedure again and compute them for $t_0 + 2\Delta t$.

Then do it again and get x and $\dot{x}$ for $t_0 + 3\Delta t$.

And so on!

This kind of iterative procedure is ideal for a computer — the computer just has to do the same things over and over again to construct the approximate trajectory. And if you want your approximation to be more accurate, you just have to make Δt smaller, so that the $\cdots$ that you throw away in (12) and (14) are smaller. If you are doing this on a computer, it doesn't matter how complicated the function F is. If we have time, we will play with the *Mathematica* notebook **lecture-1-2.nb** where is process is carried out for some simple and not so simple force laws.

But sometimes, it is nice to have analytic expressions for the trajectories. You may be able to find such things if the force is a simple enough function. For the rest of this lecture and on

Thursday, we will give examples in which this can be done, both in general, using integration, and in specific cases where the integrals can be done analytically.

Meanwhile, notice that the analysis we have just given really is a "physicist's proof" of the statement that a second order differential equation has a solution that is fixed when we know q and $\dot{q}$ at some time. We have "proved" this by actually constructing the solution! This is the best kind of proof for a physicist — one that not only tells you that the solution exists, but actually shows you how to find it. Of course this would never satisfy a mathematician — too many loose ends?

But note that we can "prove" the general theorem this way — that an n-th order differential equation requires n initial conditions. Because our original equation was a second order differential equation, we had to keep track of both $x(t)$ and $\dot{x}(t)$ while we take our small time steps. In a first order differential equation, we would only have had to keep track of $x(t)$, because the differential equation would tell us $\dot{x}(t)$ directly in terms of $x(t)$. Then we need only one initial condition. In a third order differential equation, we would have to keep track of $x(t)$, $\dot{x}(t)$ and $\ddot{x}(t)$, because the differential equation only tells us the third derivative. Then we would need three initial conditions. And so on.

Forces of the form $F(t)$

The big general principle of mechanics that we have talked about today is that we need two initial conditions per degree of freedom to specify how a system moves. Staple this in your brains. This is something we will come back to and will try to understand better. But now and next time we will spend a bit of time discussing some examples of the different ways these initial conditions can appear. There are more in Chapter 2 of Dave Morin's book. Depending on how we are doing on time, I may move this discussion to the next lecture.

The simplest and least interesting example of a force law in which the trajectories can be found formally using integration is a force depending only on t, not on x or $\dot{x}$. In this case, we can simply use the fundamental theorem of integral calculus. The acceleration is the time derivative of the velocity, so we can write $F = ma$ in this case as

$$a(t) = \frac{d}{dt}v(t) = \frac{1}{m}\,F(t) \tag{16}$$

The fundamental theorem tells us that the general solution to (16) can be written as[1]

$$v(t) = \int_{t_0}^{t} \frac{1}{m}\,F(t')\,dt' + \text{constant} \tag{17}$$

There are a number of things to note about (17):

1. We can easily verify that it works, because if we differentiate the left hand side with respect to t, we just get the acceleration, $a(t)$. On the right hand side, integral calculus tells

[1]Note that t' here is a dummy variable. We will discuss dummy variables in more detail below, but there is a possible notational confusion. The symbol t' does not mean "the derivative of t" but is just a symbol for a new independent variable. We could have called it s instead, but Dave's book uses the t' notation in this way, so I am trying to do it in lecture as well. You will have keep your wits about you to figure out what is going on, but you can almost always tell from the context. If a $'$ appears on a function, it is a derivative. If the $'$ appears on a variable, it is a new variable. Thus $f'(x)$ means the derivative with respect to x of the function $f(x)$, but $f(x')$ means the function f of the variable x' and $f'(x')$ means the derivative with respect to the variable x' of the function $f(x')$.

us that differentiating the integral with respect to t gives the integrand evaluated at t, and differentiating the constant gives zero, so we get

$$a(t) = \frac{1}{m}\,F(t) \tag{18}$$

which is $F = ma$.

2. We need to invent a new symbol for the argument of the integrand, because the previous argument only works if the time dependence appears only as the upper limit of the integral, so we have just put a prime on t. The dt' in (17) identifies the variable t' as a completely independent "dummy" variable. It is a "dummy" because in the expression (17), it doesn't have any particular value. It is just a symbol to indicate exactly what integral we are doing. This is simple but important. It has a number of consequences. Nothing depends on a dummy variable. For example in (17)

$$\int_{t_0}^{t} \frac{1}{m}\,F(t')\,dt' \tag{19}$$

we could just have well used some other symbol inside the integral:

$$\int_{t_0}^{t} \frac{1}{m}\,F(t')\,dt' = \int_{t_0}^{t} \frac{1}{m}\,F(s)\,ds = \int_{t_0}^{t} \frac{1}{m}\,F(\alpha)\,d\alpha \tag{20}$$

A particularly important consequence of this is that **you should never ever write an equation like**

$$A(t) = \int_{t_0}^{t} B(t)\,dt \tag{21}$$

This doesn't make sense because the variable t in (21) appears in two ways — as a real variable and as a dummy variable — and it can't be both. **Thus (21) doesn't mean anything at all, and if we see it on one of your problem sets or tests we will deal sternly with it!**

Also, the integral

$$\int_{t_0}^{t} B(t)\,dt' \tag{22}$$

is not the same as

$$\int_{t_0}^{t} B(t')\,dt'\,. \tag{23}$$

Both of these expressions are sensible, they are just not the same. In fact, we can do the first integral explicity because the integrand $B(t)$ doesn't depend on the integration variable, the dummy variable t', we can take it outside the integral and write -

$$\int_{t_0}^{t} B(t)\,dt' = B(t)\,\int_{t_0}^{t} dt' = (t - t_0)\,B(t)\,. \tag{24}$$

3. The integral needs some lower limit, and we have called this t_0, expecting it to have something to do with the initial conditions.

4. Finally, (17) appears to depend on both t_0 and the unknown constant. But we can determine the constant by just setting $t = t_0$, which gives

$$v(t_0) = \int_{t_0}^{t_0} \frac{1}{m} F(t')\, dt' + \text{constant} \tag{25}$$

The integral vanishes here because the range of integration is zero, so we know that

$$\text{constant} = v(t_0) = \dot{x}(t_0)\,. \tag{26}$$

We sometimes call the initial condition $\dot{x}(t_0)$ an integration constant because of the way it appears in (25).

Putting (26) back into (17) gives the final result for the velocity,

$$v(t) = \int_{t_0}^{t} \frac{1}{m} F(t')\, dt' + v(t_0) \tag{27}$$

Another way to think about the result (27) is from the following chain (really just going backwards through the same argument):

$$\int_{t_0}^{t} \frac{1}{m} F(t')\, dt' = \int_{t_0}^{t} a(t')\, dt' = \int_{t_0}^{t} \left(\frac{d}{dt'} v(t') \right) dt' = v(t) - v(t_0) \tag{28}$$

which is equivalent to (27).

Now that we know $v(t)$ as a function of t (at least formally, in terms of an integral), we can find $x(t)$ by just repeating the procedure. But v is the time derivative of x, we can write

$$x(t) = \int_{t_0}^{t} v(t')\, dt' + x(t_0) \tag{29}$$

where again, the initial condition appears as an integration constant. Then we can put (27) into (29) to get

$$x(t) = \int_{t_0}^{t} \overbrace{\left[\int_{t_0}^{t'} \frac{1}{m} F(t'')\, dt'' + v(t_0) \right]}^{v(t')} dt' + x(t_0) \tag{30}$$

Note the new variable of integration, again to keep from getting the integration variable confused with the range of integration. Finally, we can do the integral for the term proportional to $v(t_0)$, because it is just a constant, and get the result

$$x(t) = \int_{t_0}^{t} \int_{t_0}^{t'} \frac{1}{m} F(t'')\, dt''\, dt' + v(t_0)(t - t_0) + x(t_0) \tag{31}$$

The terms in (31) are easy to understand. The last term is where the object started — the initial condition for the position. The middle term describes motion with constant velocity $v(t_0)$, which is what the object would be doing if there were no force on it. And the first term is the effect of the force. Notice that (31) is valid for a constant force $F(t) = F_0$ (a constant is just a particularly simple function of t). And sure enough, we can do the integral in that case and get (11).

$$\begin{aligned} x(t) &= \int_{t_0}^{t} \int_{t_0}^{t'} \frac{1}{m} F(t'')\, dt''\, dt' + v(t_0)(t - t_0) + x(t_0) \\ &= \int_{t_0}^{t} \int_{t_0}^{t'} \frac{F_0}{m}\, dt''\, dt' + v(t_0)(t - t_0) + x(t_0) \\ &= \int_{t_0}^{t} \frac{F_0}{m}\, (t' - t_0)\, dt' + v(t_0)(t - t_0) + x(t_0) \\ &= \frac{F_0}{2m}\, (t - t_0)^2 + v(t_0)(t - t_0) + x(t_0) \end{aligned} \tag{32}$$

A $\mathcal{A}$ Implies $F = ma$

I am not going to talk about this in class, but here is a little mathematical appendix for those of you who are interested. At least for a single degree of freedom, assumption $\mathcal{A}$ actually implies, under very mild assumptions, the existence of a second order differential equation like Newton's second law for the trajectories.

To see this, let us restate $\mathcal{A}$ in mathematics. Suppose that a system has one degree of freedom so that the configuration of the system is given by the value of a single variable x (it won't matter, but it makes the language simpler to talk about only a single coordinate). Since $\mathcal{A}$ is the statement that the trajectory is determined by x and $\dot{x}$ at some given time, call the "given time" t_0, then $\mathcal{A}$ implies that we can write any possible trajectory $x(t)$ as a function of four variables: t (of course); the time t_0 at which we specify x and $\dot{x}$; and of the "initial" values $x(t_0) = x_0$ and $\dot{x}(t_0) = v_0$. $\mathcal{A}$ is just the statement that such a function exists. Since it exists, we can give it a name G and write the statement $\mathcal{A}$ in mathematics as

$$\mathcal{A} \Leftrightarrow x(t) = G\Big(x_0, v_0, t_0, t\Big) \tag{33}$$

Furthermore, because we can specify x and $\dot{x}$ at any time we choose. We can use any t_0 in (33) and get the same trajectory so long as we take x_0 and v_0 to be the values of the position and the velocity on that trajectory at the time t_0. This function G contains everything about all possible motions of the system!

The mild assumption we need to get $F = ma$ from this is a smoothness assumption. We must assume we can differentiate our function $x(t)$ with respect to time. Obviously, if we can't do this, we have no hope of getting to Newton. This assumption is really very mild.

Now here is the basic idea. Because of the meaning of the initial conditions, we can put $x_0 = x(t_0)$ and $v_0 = \dot{x}(t_0)$ into (33) and rewrite it as

$$x(t) = G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big) \tag{34}$$

This is a peculiar equation, because the trajectory $x(t)$ appears on both sides, on one side inside the function G. Because of this, the equation is telling us something about the function x that describes the trajectory. It is still more complicated than it has to be though, because it involves two different times, t and t_0. But the other peculiar thing is that the right hand side of (34) must be completely independent of t_0, because t_0 doesn't appear on the left had side at all. We can use this to get information on the trajectory that only depends on a single time. In particular, now we can start differentiating both sides of the equation with respect to t and afterwards, setting t_0 equal to t to get an equation that only depends on the the trajectory and its derivatives at a single time. We will see that differentiating zero or one time doesn't give us anything interesting - just tautologies. But differentiating twice is interesting.

So here goes. First we set $t_0 \to t$ in (34).

$$x(t) = G\Big(x(t), \dot{x}(t), t, t\Big) \tag{35}$$

This is a tautology, because we can use the time t as the initial time to specify the trajectory, and this just says that the particle is where it is supposed to be at time t. To put it another way, if I change all the ts to t_0 in (35),

$$x(t_0) = G\Big(x(t_0), \dot{x}(t_0), t_0, t_0\Big) = G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big)\Big|_{t=t_0} = x(t)|_{t=t_0} \tag{36}$$

this is just the statement of the initial condition for $x(t_0)$.

Likewise, differentiating once with respect to t gives us nothing new.

$$\dot{x}(t) = \frac{\partial}{\partial t} G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big) = \frac{\partial}{\partial t}\, G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big)\Big|_{t_0=t} \tag{37}$$

Again we can set $t_0 \to t$ because the left hand side doesn't depend on t_0. The equation (37) just says that the particle is going as fast as it should be going at time t.

$$\dot{x}(t_0) = \frac{\partial}{\partial t} G\Big(x(t_0), \dot{x}(t_0), t_0, t_0\Big) = \frac{\partial}{\partial t}\, G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big)\Big|_{t=t_0} = \frac{\partial}{\partial t}\, x(t)|_{t=t_0} \tag{38}$$

But differentiating twice with respect to t gives us a second order differential equation that the trajectory must satisfy. And this is Newton's second law.

$$\ddot{x}(t) = \frac{\partial^2}{\partial t^2} G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big) = \frac{\partial^2}{\partial t^2}\, G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big)\Big|_{t_0=t} \tag{39}$$

The right hand side of (39) is just some function of $x(t)$, $\dot{x}(t)$ and t. Thus as promised, is a second order differential equation for $x(t)$. This is just $F = ma$ with

$$F(x(t), \dot{x}(t), t) = m\, \frac{\partial^2}{\partial t^2}\, G\Big(x(t_0), \dot{x}(t_0), t_0, t\Big)\Big|_{t_0=t} \tag{40}$$

From the function that is the mathematical description of the assumption $\mathcal{A}$ that the trajectory is determined by an initial position and velocity, we can find the force that makes $F = ma$ work.

lecture 2

Topics:

Where are we?
Forces of the form $F(v)$
Example: $F(v) = -m\,\Gamma\,v$
Another example: $F(v) = -m\,\beta\,v\,|v|$
Review of the harmonic oscillator
Linearity and Time Translation Invariance
Back to $F(v) = -m\,\Gamma\,v$
Appendix - cross products - preview

Where are we?

Last time, we discussed Newton's second law — $F = ma$. I tried to convince you that this is essentially equivalent to the statement that the motion of any given classical mechanical system is determined by a set of **initial conditions**, the values of the coordinates which specify its configuration, and their first derivatives at any given time. I also suggested that this is a very deep and interesting fact about the world, and promised that we would come back to it at the end of the course and give at least a provisional explanation of it.

We also discussed how to solve for the motion of a system numerically by keeping track of q and $\dot{q}$ as functions of time, and using the Taylor expansion and $F = ma$ to calculate approximately how they change in a small time step Δt. By putting together many small time steps, we can trace out the trajectory of the system. This procedure works for any any number of degrees of freedom, and it should convince you that in principle, giving a second order differential equation for the configuration of a classical does just what expect - it determines the trajectory in terms a set of two initial conditions per degree of freedom.

In a sense, our numerical analysis completely solves the problem - at least least your computer can construct the solution to any problem. But for us people, it is nice to have analytic solutions that we can use to develop our intuition. So we also talked about systems with one degree of freedom in which the force depends only on time. In this VERY special case, we found that we could write down the formal solution simply by integration. Then if the integral can be done analytically, we get a completely analytic solution. We will go over this again quickly in lecture because we didn't get to talk about it last time.

Today, we will give some more examples of very special systems in which we can do more than just solving numerically. In some sense, I will just be showing you a collection of dirty tricks, because it is only in very special cases that they work. But more generally, today's lecture should be regarded as a bunch of examples of the different ways in which initial conditions can enter into the solutions of classical mechanics problems. We always need two initial conditions per degree of freedom. But they appear in the actual trajectories in many different ways. In fact, I have something else in mind as well. At the end, when we come to discuss the harmonic oscillator, we will see that there are some very important general principles at work. These will be very useful, and we will come back to them many times.

Forces of the form $F(v)$

Formally, it is almost as easy to solve for the trajectory for a system of one degree of freedom in which the force depends only on v as for a force that depends only on t. For a force $F(v)$ depending only on v, we write $F = ma$ as

$$m\,\frac{dv}{dt} = F(v) \tag{1}$$

We can rewrite this as

$$dt = \frac{m}{F(v)}\,\frac{dv}{dt}\,dt \tag{2}$$

which we can simplify just by integration. Again we have to remember that the integration variables must be distinguished from the ranges of integration. Again I will use the notation in Morin's book and just put primes on them. Putting in all the steps very formally, we get

$$\int_{t_0}^{t} dt' = t - t_0 = \int_{t_0}^{t} \frac{m}{F(v')}\,\frac{dv'}{dt'}\,dt' = \int_{v(t_0)}^{v(t)} \frac{m}{F(v')}\,dv' \tag{3}$$

The last step is a change of variable in the integration from t' to $v' = v(t')$. A quick and dirty way of getting to (3) is to write

$$dt = \frac{m}{F(v)}\,dv \tag{4}$$

and integrate both sides.[1] This is a perfectly good way of looking at it. The equation (4) describes how a small change in t is related to the small change in v that takes place during the small change in t. Integration adds the small changes up.

Equation (3) implicitly determines $v(t)$ in terms of $t - t_0$ and the initial condition $v(t_0)$. And once we know $v(t)$, we can integrate to get the trajectory $x(t)$ (using the initial condition $x(t_0)$ as usual).

Let's see how this works in an important example, describing a frictional force.

Example: $F(v) = -\beta\,v = -m\,\Gamma\,v$

If an object moves very slowly through a thick gas or a liquid, there is almost always a frictional force approximately proportional to the velocity (and in the opposite direction). What "slowly" means depends on the size of the object and the viscosity of the stuff it is moving through. We will come back later in the lecture to why this force should exist. For now, let's see how (3) works for such force, which we can take to have the form[2]

$$F(v) = -\beta\,v = -m\,\Gamma\,v \tag{5}$$

(3) becomes

$$\int_{t_0}^{t} dt' = -\int_{v(t_0)}^{v(t)} \frac{1}{\Gamma\,v'}\,dv' \tag{6}$$

[1] When we do that, we must turn them into dummy variables, which we do, as usual, by putting primes on the.

[2] The factor of m is just there to make the units of the constant Γ simpler. It makes the ms cancel in (3). Physically, the factor doesn't really make any sense, because the frictional force on an object just depends on its size, shape and surface properties — not on its mass. But of course the essential physics doesn't change when you leave out the m and write things in terms of β - the formulas just look a bit more complicated.

This looks nicer if we put the Γ on the other side

$$\Gamma \int_{t_0}^{t} dt' = -\int_{v(t_0)}^{v(t)} \frac{1}{v'}\, dv' \tag{7}$$

or

$$\Gamma\,(t-t_0) = -\ln v(t) + \ln v(t_0) \tag{8}$$

or

$$\ln v(t) = \ln v(t_0) - \Gamma\,(t-t_0) \tag{9}$$

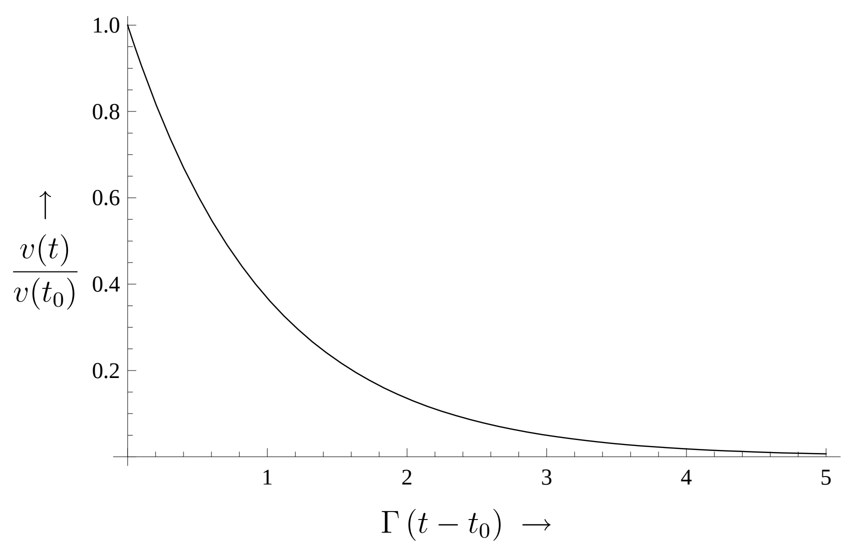
or

$$e^{\ln v(t)} = e^{\ln v(t_0) - \Gamma\,(t-t_0)} \tag{10}$$

or

$$v(t) = v(t_0)\, e^{-\Gamma\,(t-t_0)} \tag{11}$$

This makes sense. The constant Γ has units of $1/T$, so the exponent is dimensionless, as it should be. Thus the units work in (11). We can also check this in the limit $\Gamma \to 0$. In this limit, where the frictional force disappears, the velocity goes to a constant, as it should. For nonzero Γ, the frictional force causes the velocity to gradually drop off. Because the force is proportional to v, the drop-off gets slower and slower as the velocity gets smaller, in this case, exponentially. A plot of the velocity (in units of v_0) as a function of time (in units of $1/\Gamma$) has an exponential shape that I hope is familiar:



Now let's compute $x(t)$ by integrating (11).

$$x(t) = x(t_0) + \int_{t_0}^{t} v(t')\, dt' = x(t_0) + \int_{t_0}^{t} v(t_0)\, e^{-\Gamma\,(t'-t_0)}\, dt' \tag{12}$$

Doing the integral is easier if we substitute

$$u = e^{-\Gamma\,(t'-t_0)} \qquad du = -\Gamma\, e^{-\Gamma\,(t'-t_0)}\, dt' \tag{13}$$

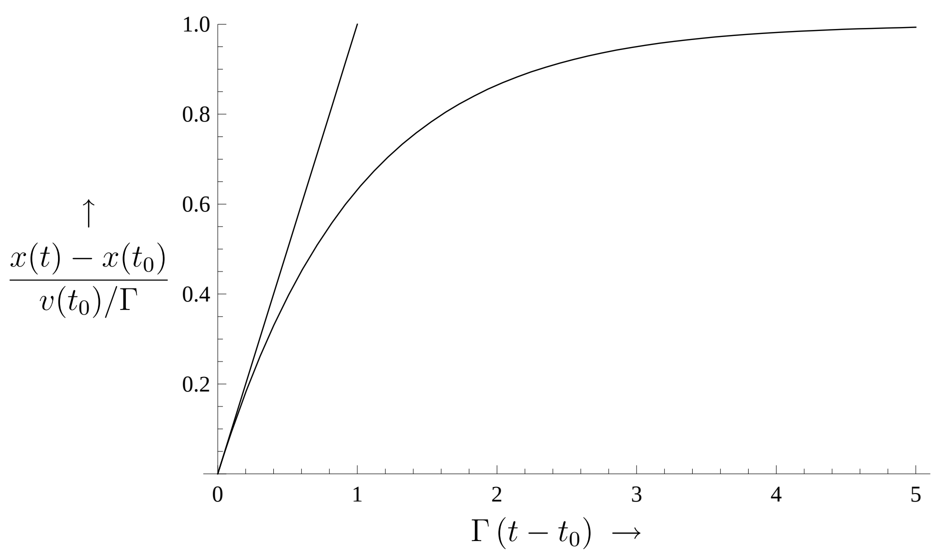
we get

$$x(t) = x(t_0) - \frac{v(t_0)}{\Gamma} \int_1^{e^{-\Gamma(t-t_0)}} du \tag{14}$$

which gives

$$x(t) = x(t_0) + \frac{v(t_0)}{\Gamma} \left(1 - e^{-\Gamma(t-t_0)}\right) \tag{15}$$

Note that the object moves a finite distance, $v(t_0)/\Gamma$, in infinite time. The distance traveled looks like this, where for comparison, I have included the linear extrapolation of the initial velocity.

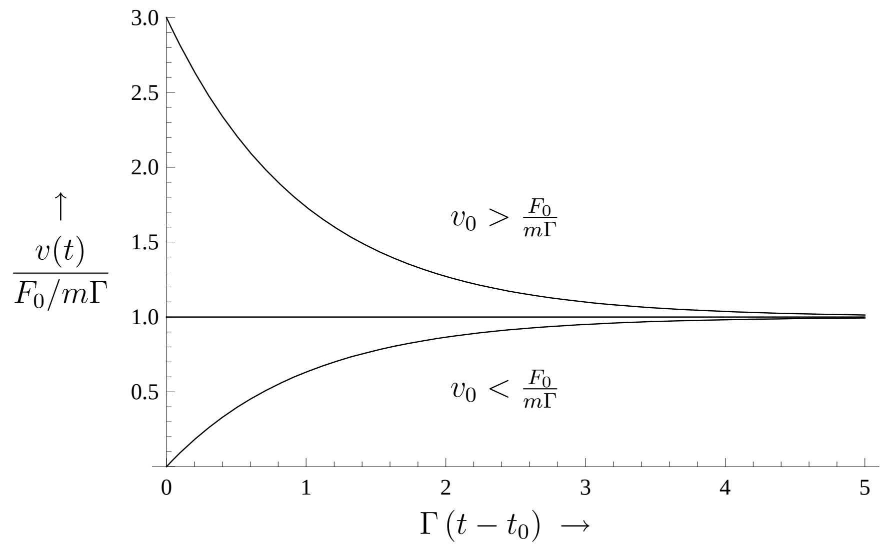


As usual, it is really really fun to play with this in with *Mathematica*. Check out **lecture-2-1.nb** included with the lecture notes.

In Dave Morin's book, he works out a slightly more interesting example in which there is also a constant component of the force. This is relevant to objects falling in a gravitational field and also subject to friction. If the force looks like $F(v) = F_0 - m\,\Gamma\,v$, the velocity as a function of time is

$$v(t) = v(t_0)\,e^{-\Gamma(t-t_0)} + \frac{F_0}{m\,\Gamma}\left(1 - e^{-\Gamma(t-t_0)}\right) \tag{16}$$

One interesting thing about this is that $v(\infty) = \frac{F_0}{m\,\Gamma}$ independent of v_0. Whatever the initial conditions, v approaches the same "terminal" velocity, as illustrated in the graph below:



If Γ is very large, this gives a kind of Aristotelian physics. **When you push on something, it moves with a velocity proportional to the force.** If Γ is large enough, you might not notice that it takes a little time ($1/\Gamma$) to settle down to this terminal velocity.

Another example: $F(v) = -m\,\beta\,v\,|v|$

If an object moves rapidly through a gas, there is a frictional force proportional to $v\,|v|$. The funny form of the force with absolute value, $|v|$, is something that you will explore on the problem set, so I won't talk about it now. I will just assume that $v > 0$ so that $v\,|v| \rightarrow v^2$. The force arises because the object knocks the gas molecules out of the way, and the force is proportional both to momentum change of the knocked molecules (which contributes a factor of $-v$ to the force) and to the number of molecules that are knocked per unit time (a factor of $|v|$). Now let's begin the same analysis for such a force, of the form

$$F(v) = -m\,\beta\,v^2 \tag{17}$$

Now (3) becomes

$$\int_{t_0}^{t} dt' = -\int_{v(t_0)}^{v(t)} \frac{1}{\beta\,v'^{\,2}}\,dv' \tag{18}$$

This looks nicer if we put the β on the other side

$$\beta \int_{t_0}^{t} dt' = -\int_{v(t_0)}^{v(t)} \frac{1}{v'^{\,2}}\,dv' \tag{19}$$

or

$$\beta\,(t-t_0) = \frac{1}{v(t)} - \frac{1}{v(t_0)} \tag{20}$$

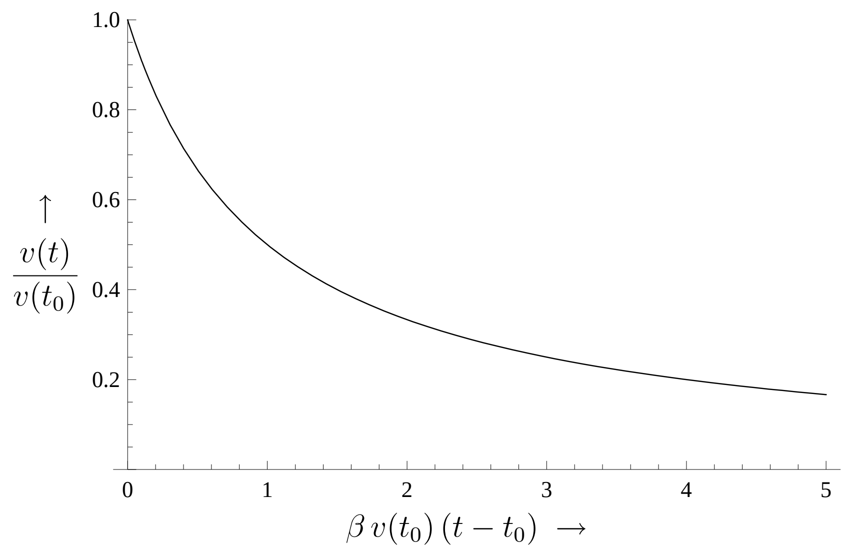
or

$$\frac{1}{v(t)} = \frac{1}{v(t_0)} + \beta\,(t-t_0) = \frac{1}{v(t_0)} + \frac{\beta\,v(t_0)\,(t-t_0)}{v(t_0)} = \frac{1+\beta\,v(t_0)\,(t-t_0)}{v(t_0)} \tag{21}$$

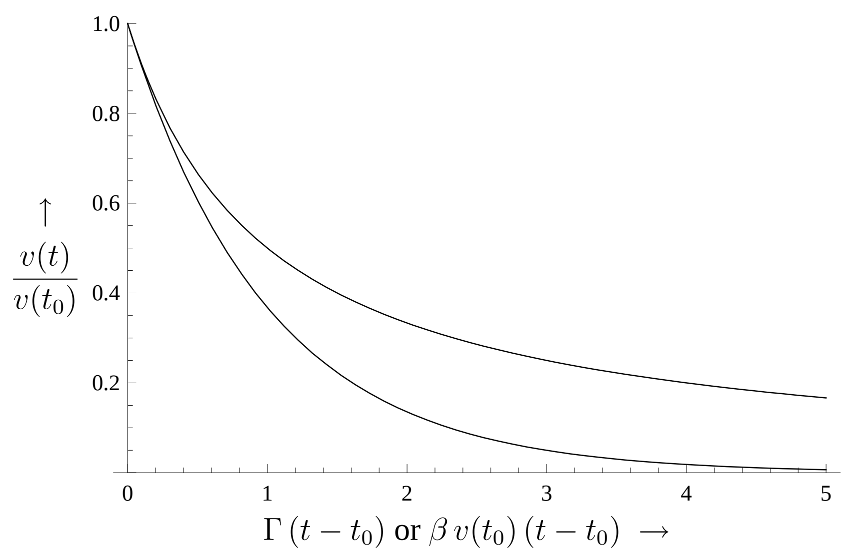
or

$$v(t) = \frac{v(t_0)}{1 + \beta\, v(t_0)\, (t - t_0)} \tag{22}$$

This looks a bit different from the exponential.



Compared to an exponential that initially falls at the same rate, this one falls off more slowly later, because the the frictional force gets smaller faster as the velocity decreases.



I won't work out the details here, but you can now integrate this again to get the position as a function of time:

$$x(t) = x(t_0) + \frac{1}{\beta} \log\Big(1 + \beta\, v(t_0)\, (t - t_0)\Big) \tag{23}$$

Again, we can use **lecture-2-1.nb** on the *Mathematica* page to see what this looks like.

I did another example basically just to show how differently the initial condition (the value of $v(t_0)$ in this case) comes into to the final result in different cases [compare (22) and (23) with

(11) and (15)]. The trajectories always depend on two initial conditions, but what that dependence looks like depends on the form of the force. You have to think and keep your wits about you — if you try to remember formulas for all the different possible force laws, you will just get confused. Instead concentrate on understanding the derivations! But nevertheless the matra — **two initial conditions per degree of freedom** will be a useful check that you are doing sensible things.

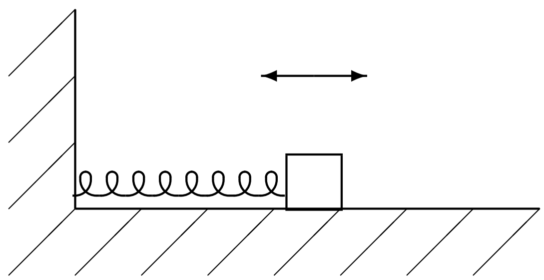
Review of the harmonic oscillator

Dave Morin talks about forces of the form $F(x)$ in his chapter 3, which are reading. But I am going to postpone a general discussion of until after the problem set. The general treatment of these forces is essentially just an application of the principle of conservation of energy, so it makes sense to wait until we talk about energy. Instead, we will review the simplest example of such a force - the harmonic oscillator with

$$F(x) = -K\,x\,. \tag{24}$$

There is another good reason for thinking about this. In general, if the force depends on both x and v (or x, v and t), things are much harder. But by cleverly adding damping and driving to the harmonic oscillator, (24), we can include both x and v dependence, and t dependence as well, and still say interesting things rather simply. We will talk about this starting next time.

The harmonic oscillator is a very important mechanical system. We will start today at the beginning but will go on to try to explain why it is important. The simplest example of a harmonic oscillator is a mass attached to a spring, shown here sitting on a surface, which I will assume is frictionless.



(25)

There is an equilibrium position in which the spring is neither stretched nor compressed. In equilibrium, the spring produces no force on the mass. If the mass moves so that the spring is not in equilibrium, the spring exerts a force on the mass that tends to move it back to equilibrium. If the mass is released from out of equilibrium, it accelerates towards the equilibrium position. When it gets to the equilibrium position, the force goes to zero, but it is moving, so inertia keeps it going to the other side of equilibrium. This is the classic recipe for oscillation.

But so far, this is just oscillation. What makes this a harmonic oscillator is the special force law of the ideal "Hooke's Law" spring, in which the force is proportional to minus the displacement from equilibrium. If the block moves in the x direction, and we choose a coordinate system in which $x = 0$ at equilibrium, then the force on the block looks like

$$F = -K\,x \tag{26}$$

where K is called the spring constant. The larger the spring constant, the stiffer the spring.

Now Newton says

$$F = m\,a = m\,\frac{d^2}{dt^2}x = m\,\ddot{x} = -K\,x \tag{27}$$

You probably know that the general solution to the second order differential equation (27) is

$$x(t) = a \cos \omega t + b \sin \omega t \tag{28}$$

where

$$\omega \equiv \sqrt{\frac{K}{m}} \tag{29}$$

If you don't know this by heart, that is fine. Next time we will learn how to solve such things in general in a very easy way. But meanwhile it is always possible to check that our solution works even if we don't know where it comes from. Differentiating (28) twice gives

$$\dot{x}(t) = -a\omega \sin \omega t + b\omega \cos \omega t \tag{30}$$

$$\ddot{x}(t) = -a\omega^2 \cos \omega t - b\omega^2 \sin \omega t = -\omega^2 (a \cos \omega t + b \sin \omega t) = -\omega^2 \, x(t) \tag{31}$$

so that

$$m \, \ddot{x}(t) = -m \, \omega^2 \, x(t) = -m \, \frac{K}{m} \, x(t) = -K \, x(t) \tag{32}$$

The constant ω is called the **angular frequency**. It is fixed by the physics - the values of the mass and the spring constant. But as we expected because we need two initial conditions/degree of freedom, there are an infinite number of possible trajectories, labeled by two constants, here a and b. Again as expected, we can determine a and b by imposing initial conditions. The details of the way the constants enter are different here than in our previous examples, but at least the number of initial conditions is right — two constants for one degree of freedom. Setting $t = 0$ in (28) gives

$$x(0) = a \cos 0 + b \sin 0 = a \tag{33}$$

Thus a is the position of the mass at $t = 0$. Setting $t = 0$ in (30) gives

$$v(0) = \dot{x}(0) = -a\omega \sin 0 + b\omega \cos 0 = b\omega \tag{34}$$

Thus $b\omega$ is the velocity of the mass at $t = 0$. Thus we can rewrite (28) as

$$x(t) = x(0) \cos \omega t + \frac{v(0)}{\omega} \sin \omega t \tag{35}$$

As we will discuss in more detail later, we can easily adapt (35) to use the initial conditions at any time $t = t_0$ as follows:

$$x(t) = x(t_0) \cos\Big[\omega(t - t_0)\Big] + \frac{v(t_0)}{\omega} \sin\Big[\omega(t - t_0)\Big] \tag{36}$$

This comes alive in **lecture-2-2.nb** on the *Mathematica* page.

It is interesting to collect the various examples we have discussed in one table.

Force	trajectory $x(t)$
$F_0 \Rightarrow$	$x(t_0) + v(t_0)\,(t-t_0) + \frac{F_0}{2m}(t-t_0)^2$
$-m\,\Gamma\,v \Rightarrow$	$x(t_0) + \frac{v(t_0)}{\Gamma}\,\left(1-e^{-\Gamma\,(t-t_0)}\right)$
$-m\,\beta\,v^2 \Rightarrow$	$x(t_0) + \frac{1}{\beta}\log\Big(1+\beta\,v(t_0)\,(t-t_0)\Big)$
$-m\,\omega^2\,x \Rightarrow$	$x(t_0)\cos\Big[\omega(t-t_0)\Big] + \frac{v(t_0)}{\omega}\sin\Big[\omega(t-t_0)\Big]$

(37)

What I think is interesting about this is that it shows very clearly that while each of these force laws requires two initial conditions, which we can take to be the position and velocity at the time t_0, the way in which these two constants appear differs dramatically from one force law to another. This is an important lesson. There are always two initial conditions per degree of freedom, the way the initial conditions appear in the trajectory depends on the force law.

There is a very different way of writing the solution (28). Using (I hope familiar) trigonometric identities, we can write, with

$$t_0 = 0 \qquad a \equiv x(0) \qquad b \equiv v(0)/\omega \tag{38}$$

$$x(t) = a\cos\omega t + b\sin\omega t = c\cos(\omega t - \phi) \tag{39}$$

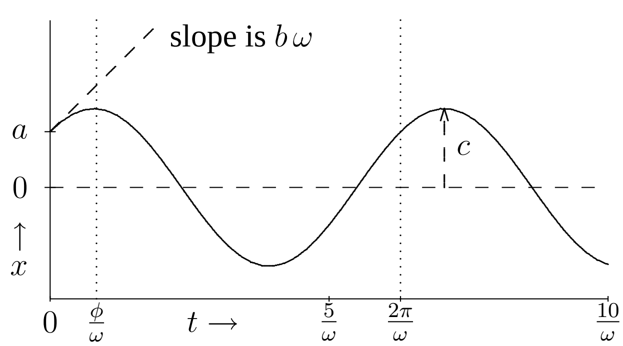
$$= c\,(\cos\omega t\cos\phi + \sin\omega t\sin\phi) \tag{40}$$

$$= c\cos\phi\cos\omega t + c\sin\phi\sin\omega t \tag{41}$$

$$a = c\cos\phi \qquad b = c\sin\phi \tag{42}$$

$$c = \sqrt{a^2+b^2} \qquad \phi = \arctan\frac{b}{a} \tag{43}$$

The constant c is called the **amplitude** of the motion and ϕ is the phase of the oscillation. (40)-(43) show that the two constants c and ϕ carry the same information as the two constants a and b. Both pairs are determined by the initial conditions and either pair completely determines the trajectory. The connection between these constants and the trajectory is illustrated in (44), a graph of the trajectory $x(t)$ versus t.



(44)

Because the $\sin$ and $\cos$ functions are periodic with period 2π,

$$x(t+2\pi/\omega) = c\cos\left[\omega(t+2\pi/\omega)-\phi\right] = c\cos(\omega t-\phi+2\pi) = c\cos(\omega t-\phi) = x(t) \tag{45}$$

and thus the motion repeats after a time

$$\tau = 2\pi/\omega \quad \text{which is called “the period of the oscillation.”} \tag{46}$$

The frequency (as opposed to the angular frequency) is

$$\nu = \frac{\omega}{2\pi} = \frac{1}{\tau} \tag{47}$$

The product ωt — the argument of $\sin$ and $\cos$ is an angle.

$$\text{Note that} \quad \boxed{\nu \text{ has units } \frac{\text{cycles}}{\text{sec}}} \quad \text{and} \quad \boxed{\omega \text{ has units } \frac{\text{radians}}{\text{sec}}}. \tag{48}$$

These are really the same units in the sense of dimensional analysis because 2π is dimensionless, but thinking about the units as stated in (48) will help you remember the factor of 2π which is just the conversion factor between radians and cycles. It will make even more sense soon when we make the connection between harmonic oscillation and uniform circular motion.

So what is it that is special about harmonic oscillation? — two general principles —

Linearity and Time Translation Invariance

Harmonic oscillation occurs almost everywhere. There are lots of physical systems whose motion is described, at least approximately, by solutions to the same equation as the mass on a spring. Why is that?

$$m\,\frac{d^2}{dt^2}x + K\,x = 0 \tag{49}$$

There are two key features.

1: Time translation invariance — there is no explicit dependence on time. There are dts — in the derivatives with respect to time. But there are no ts. When the equation for motion has this property, it follows that **if $x(t)$ is a solution then so is $x(t+a)$.** You can easily prove this using the chain rule.

Time translation invariance is particularly useful for the harmonic oscillator because of the other special property of its force law.

2: Linearity — because all the terms are proportional to one power of x or its derivatives, we can make new solutions as linear combinations of old ones — **if $x_1(t)$ and $x_2(t)$ are solutions then so is $A\,x_1(t)+B\,x_2(t)$.** Again, this can be checked explicitly.

Time translation invariance is physically very reasonable. The laws of physics don't change with time. Or if they do, they do so very slowly (maybe on the scale of the age of the universe, 15 billion years or so), so that we can't tell. There are lots of systems in which the coefficients (things like mass, spring constant, inductance, capacitance, etc) are determined directly by the laws of physics and are therefore independent of time.

Linearity may seem more like an obscure mathematical concept — what does it have to do with physics? Why should there be so many systems that are approximately linear? The word "approximately" here is important. Probably, linearity is never exactly true for a classical system, but it is often an excellent approximation.

So suppose that you are studying oscillations in some system, and all you know about it is that it oscillates about a point of equilibrium. Let's call the quantity that measures the displacement from equilibrium x, to make it look like the mass on a spring. We expect that the time evolution of the system can be described by the solutions to some second order differential equation (like $F = m\,a$):

$$\frac{d^2}{dt^2}x = \mathcal{F}(x) \tag{50}$$

where $\mathcal{F}$ is the analog of the F/m in $F = m\,a$. Because I've assumed that we are in equilibrium at $x = 0$, it must be that

$$\mathcal{F}(0) = 0 \tag{51}$$

Now, assuming that $\mathcal{F}(x)$ is smooth, we can use the most important formula in physics — and you can probably guess what that is — the Taylor expansion

$$\mathcal{F}(x) = \mathcal{F}(0) + x\mathcal{F}'(0) + \frac{1}{2}x^2\mathcal{F}''(0) + \cdots = x\mathcal{F}'(0) + \frac{1}{2}x^2\mathcal{F}''(0) + \cdots \tag{52}$$

Then unless $\mathcal{F}'(0)$ is exactly zero, the first term will dominate for sufficiently small x. This is why linearity is so important. Most functions in physics are smooth. Most of the time, there is no particular reason for $\mathcal{F}'(0)$ to be zero, so it isn't. Thus the equations of motion for most systems are linear for sufficiently small x.

I should perhaps just note that the most important example of linearity has little to do with classical mechanics — it is quantum mechanics itself. Just as classical sound waves or electromagnetic waves can add together and sometimes interfere constructively and sometimes destructively, so also the mysterious quantum matter waves that describe quantum states can be added. As far as we know, the linearity of quantum mechanics is exact, not an approximation. At least, very sensitive experiments have failed to find any nonlinearities.

Next time, we will discuss the consequences of time translation invariance and linearity in detail. We will see that when the physics of a system obeys these two general principles, the trajectories can be written in a very simple form, as sums of exponentials. But sometimes, the exponentials will be complex — that is they will involve $i = \sqrt{-1}$. Furthermore the generality of argument will allow us to extend this result to systems with arbitrary numbers of degrees of freedom, with rather dramatic results.

Back to $F(v) = -m\,\Gamma\,v$

For now, let us now return to the force law that we considered at the beginning of this lecture, $F(v) = -m\,\Gamma\,v$. What is the physics of this force law? We saw that when friction arises because stuff gets knocked out of the way, we get a v^2 dependence on velocity. But this v^2 dependence is not linear. When v^2 is sufficiently small, we might expect that this effect will become negligible compared to other effects that give a linear dependence on velocity. Indeed, in most liquids, we get an approximately linear dependence of the frictional force on velocity for objects that are moving slowing enough. Then the molecules of the liquid are not so much knocked as they are gently pushed out of the way, so that the process is very smooth and reversible. This is a good excuse to show you one of my very favorite demos, illustrating the smoothness that one gets in a very viscous medium, in which the linear regime is easy to reach.

Appendix - cross products - preview

The cross product is essentially just an antisymmetric combination of two vectors. This antisymmetric combination of two vectors in interesting because it defines a plane, and planes are intimately connected with rotations. The particularly convenient thing about this combination in three dimensional space is that it behaves like another vector. The cross product is the mathematical statement of the fact the antisymmetric combination of two vectors in three dimensional space defines a plane which in turn defines another vector. The geometrical definition of the cross product is a good way to see that it behaves like a vector under rotations, so we will start with that. Then I will indicate how we can show that this geometrical definition is equivalent to a definition given in terms of components.

The geometrical definition is this:

> Given two vectors, $\vec{A}$ and $\vec{B}$, the object $\vec{A} \times \vec{B}$ is a vector with magnitude $\left|\vec{A}\right| \left|\vec{B}\right| \sin\theta$ where θ is the angle between $\vec{A}$ and $\vec{B}$ defined as a positive angle between 0 and π. The direction of $\vec{A} \times \vec{B}$ is perpendicular to the plane formed by $\vec{A}$ and $\vec{B}$ with the sign determined by the right-hand rule. (53)

With this definition, it is easy to understand why $\vec{A} \times \vec{B}$ behaves like a vector under rotations. The magnitude doesn't change under a rotation because $\left|\vec{A}\right|$, $\left|\vec{B}\right|$ and $\sin\theta$ are all unchanged. And the direction rotates properly because it is tied to the directions of $\vec{A}$ and $\vec{B}$.

It is crucial that the cross product $\vec{A} \times \vec{B}$ is antisymmetric in the two vectors $\vec{A}$ and $\vec{B}$,

$$\vec{A} \times \vec{B} = -\vec{B} \times \vec{A} \tag{54}$$

In the geometrical definition, this follows from the application of the right hand rule. If you interchange $\vec{A}$ and $\vec{B}$, the cross product changes direction because the right hand rule goes from $\vec{B}$ to $\vec{A}$ rather than from $\vec{A}$ to $\vec{B}$. This antisymmetry ensures that either the two vectors $\vec{A}$ and $\vec{B}$ define a plane or the antisymmetric combination vanishes. Then the fact that in three dimensional space, there is a unique direction perpendicular to a given plane allows us to turn the antisymmetric combination into a vector.

The geometrical definition, (53), is equivalent to the following component definition,

$$\left[\vec{A} \times \vec{B}\right]_x = A_y B_z - A_z B_y \,, \quad \left[\vec{A} \times \vec{B}\right]_y = A_z B_x - A_x B_z \,, \quad \left[\vec{A} \times \vec{B}\right]_z = A_x B_y - A_y B_x \,, \tag{55}$$

where we are using a notation for vector components in which

$$\left[\vec{A}\right]_x = A_x \,, \quad \left[\vec{A}\right]_y = A_y \,, \quad \left[\vec{A}\right]_z = A_z \,. \tag{56}$$

If you have not seen cross products before in your math courses, you can look at the demonstration of this equivalence below. We will be using cross products a lot later on in the course when we talk about rotations in three dimensions.

To prove that (53) and (55) are equivalent, we first show that

$$\left(\vec{A} \times \vec{B}\right) \cdot \vec{A} = \left(\vec{A} \times \vec{B}\right) \cdot \vec{B} = 0 \tag{57}$$

We can do this by explicit calculation. For example,

$$\left(\vec{A} \times \vec{B}\right) \cdot \vec{A} = (A_y B_z - A_z B_y)A_x + (A_z B_x - A_x B_z)A_y + (A_x B_y - A_y B_x)A_z = 0 \tag{58}$$

The calculation for $\vec{B}$ is similar (as it must be because of the antisymmetry of the cross product - (58) just says that the dot product of the cross product with the first vector in the cross product vanishes - and because of antisymmetry the same must be true for the second vector in the cross product). Thus $\vec{A} \times \vec{B}$ is perpendicular to both A and B and therefore perpendicular to the plane they form, just as in the geometrical definition.

You can see that the magnitude of the object given by (55) is right by explicitly calculating its square.

$$\left(\vec{A} \times \vec{B}\right) \cdot \left(\vec{A} \times \vec{B}\right) = (A_y B_z - A_z B_y)^2 + (A_z B_x - A_x B_z)^2 + (A_x B_y - A_y B_x)^2 \tag{59}$$

$$= A_x^2 B_y^2 + A_x^2 B_z^2 + A_y^2 B_x^2 + A_y^2 B_z^2 + A_z^2 B_x^2 + A_z^2 B_y^2 - 2A_x B_x\, A_y B_y - 2A_x B_x\, A_z B_z - 2A_y B_y\, A_z B_z \tag{60}$$

If we add and subtract $A_x^2 B_x^2 + A_y^2 B_y^2 + A_z^2 B_z^2$ to this, the positive terms can be factored into

$$\left(A_x^2 + A_y^2 + A_z^2\right)\left(B_x^2 + B_y^2 + B_z^2\right) \tag{61}$$

and the negative terms into

$$-\left(A_x B_x + A_y B_y + A_z B_z\right)^2 \tag{62}$$

so we can write (59) as

$$\left(\vec{A} \times \vec{B}\right) \cdot \left(\vec{A} \times \vec{B}\right) = \left|\vec{A}\right|^2 \left|\vec{B}\right|^2 - \left(\vec{A} \cdot \vec{B}\right)^2 = \left|\vec{A}\right|^2 \left|\vec{B}\right|^2 (1 - \cos^2\theta) = \left|\vec{A}\right|^2 \left|\vec{B}\right|^2 \sin^2\theta \tag{63}$$

Finally, you can see the right-hand rule by calculating an example, like $\hat{x} \times \hat{y} = \hat{z}$. Thus we have checked that the component definition (55) is equivalent to the geometrical definition (53).

lecture 3

Topics:

Where are we?
Consequences of Time Translation Invariance and Linearity
$F = -m\Gamma v$ again
The harmonic oscillator again
The damped harmonic oscillator
Uniform circular motion in the complex plane
Harmonic oscillation for more degrees of freedom
The double pendulum

Where are we?

Last time, we saw how initial conditions appeared in a number of different examples of force laws. The last of these, the harmonic oscillator, is a particularly important system because it has two general properties, time translation invariance and linearity, that appear in many many physical systems.

Because linearity went by pretty quickly last time, let me briefly review how it works. The equation of motion, $F = ma$, for the harmonic oscillator is linear because there is a single x in each term. It can be written as

$$m \frac{d^2x}{dt^2} + K\,x = 0 \tag{1}$$

We can think of this as a single "operator" acting on x.

$$= \left(m \frac{d^2}{dt^2} + K\right)\,x = 0 \tag{2}$$

In this form, it may be more clear why you can add solutions together to get new solutions. If you have two solutions, $x_1(t)$ and $x_2(t)$, you can form an arbitrary linear combination of the two and the result will be another solution.

$$\begin{gathered} \left(m \frac{d^2}{dt^2} + K\right)\,x_1(t) = 0 \\ \left(m \frac{d^2}{dt^2} + K\right)\,x_2(t) = 0 \\ \Rightarrow \\ \left(m \frac{d^2}{dt^2} + K\right)\,\big(a\,x_1(t) + b\,x_2(t)\big) = 0 \end{gathered} \tag{3}$$

because

$$\begin{aligned}
&\left(m\,\frac{d^2}{dt^2}+K\right)\,\Big(a\,x_1(t)+b\,x_2(t)\Big) \\
&=\left(m\,\frac{d^2}{dt^2}+K\right)\,\Big(a\,x_1(t)\Big) \\
&+\left(m\,\frac{d^2}{dt^2}+K\right)\,\Big(b\,x_2(t)\Big) \\
&=a\,\left(m\,\frac{d^2}{dt^2}+K\right)\,x_1(t) \\
&+b\,\left(m\,\frac{d^2}{dt^2}+K\right)\,x_2(t) \\
&=0
\end{aligned} \tag{4}$$

This fact has remarkable consequences, as we will see shortly.

Consequences of Time Translation Invariance and Linearity

Time translation invariance is an example of a symmetry. The physics of the harmonic oscillator looks the same if all clocks are reset by the same amount. When a symmetry is combined with the property of linearity, the result is an extremely powerful tool for studying the solutions of the system's equation of motion. The reason is that because of linearity, the solutions of the equation of motion form what mathematicians call a linear space. You can add them together and multiply them by constants and you still have solutions. Because of this, we can use the tools of linear algebra to understand them. You should all be used to some of these tools even if you have not studied much linear algebra, because you all familiar with one linear space - the space of vectors in three dimensions. One of the really important things about linear spaces is that you are free to choose any convenient basis - like $\hat{x}$, $\hat{y}$ and $\hat{z}$ in three dimensions.

In the linear space of solutions of the harmonic oscillator, we can choose a convenient set of **basis** solutions that behave as simply as possible under time translations. For the symmetry of time translation, it is a mathematical fact that the simplest basis solutions are just exponentials. We can always find solutions of the form[1]

$$z(t)=z(0)\,e^{Ht} \tag{5}$$

where H is some constant. What is special about this form (and I am not going to discuss this in detail - I hope that you will see this beautiful argument in more detail in Physics 15c) is that when you change the setting of your clock by taking $t\to t+a$, the exponential (5) is the only function that just changes by a multiplicative constant,

$$z(t+a)=z(0)\,e^{H(t+a)}=z(0)\,e^{Ht}\,e^{Ha}=e^{Ha}\,z(t) \tag{6}$$

You can always use the linearity of the space of solutions to find particularly convenient solutions that behave in this simple way under time translations - and then the result has to be an exponential.

[1]This is essentially equivalent to the statement that a linear differential equation with constant coefficients always has an exponential solution — see Morin's text.

Once you realize that the solutions are just exponentials, you can find the possible values of H in a very simple way. If you do take (5) and put it into the equation of motion, time derivatives acting on e^{Ht} just bring down factors of H. This converts the differential equation into an algebraic equation.

$$\frac{d}{dt} \to H \tag{7}$$

Then once we have the basis solutions, we can form linear combinations to satisfy the initial conditions, in the same way that we can write a general vector in three dimensional space as a sum of coordinates times basis vectors

$$\vec{r} = x\,\hat{x} + y\,\hat{y} + z\,\hat{z}$$

This simple dependence of $\vec{r}$ on the coordinates x, y and z is the power of linearity at work. We will see that it works in a similarly simple way for the initial conditions in a linear and time translation invariant mechanical system.

$F = -m\Gamma v$ again

Let's do this for the linear frictional force we talked about last time, of the form

$$-m\,\Gamma\,v \tag{8}$$

with

$$\Gamma \geq 0 \tag{9}$$

We know how to solve this directly by integrating, as we did last time, but you will notice, I hope that this force law is also time translation invariant and linear. There is just a single factor of x in

$$v = \frac{dx}{dt} \tag{10}$$

and t enters only through derivatives. Therefore, we expect that the solution is a sum of exponentials in time multiplied by constant coefficients that depend on the initial conditions. The equation of motion has the form

$$m\,\ddot{x} = -m\,\Gamma\,\dot{x} \quad \text{or} \quad m\left(\ddot{x} + \Gamma\,\dot{x}\right) = 0 \tag{11}$$

This allows us to solve the problem in a different way, by assuming that the solution is a linear combination of exponentials. If we put in a trial solution of the form

$$x(t) = A\,e^{H\,t} \tag{12}$$

each dot becomes a factor of H and we get

$$m\left(H^2\,A\,e^{H\,t} + \Gamma\,H\,A\,e^{H\,t}\right) = 0 = m\,A\,e^{H\,t}\left(H^2 + \Gamma\,H\right) = 0 \tag{13}$$

The factor $m\,A\,e^{H\,t}$ never vanishes except for the uninteresting case $A = 0$, so we can just cancel it and conclude that

$$H^2 + \Gamma\,H = (H + \Gamma)\,H = 0 \Rightarrow H = -\Gamma \ \text{ or } \ H = 0 \tag{14}$$

so there are two simple solutions,

$$x(t) = A\,e^0 = A \quad \text{and} \quad x(t) = A\,e^{-\Gamma\,t} \tag{15}$$

The general solution, because of linearity, is then a combination of these two solution with coefficients that depend on the initial conditions:

$$x(t) = A + B\,e^{-\Gamma\,t} \tag{16}$$

Comparing with the solution that we got by direct integration,

$$x(t) = x(t_0) + \frac{v(t_0)}{\Gamma}\,\left(1 - e^{-\Gamma\,(t-t_0)}\right) \tag{17}$$

we see that it is equivalent to (16) with

$$A = x(t_0) + \frac{v(t_0)}{\Gamma} \tag{18}$$

and

$$B = -\frac{v(t_0)}{\Gamma}\,e^{\Gamma\,t_0} \tag{19}$$

I hope you agree that this calculation was quite a bit easier and quicker than integrating. Using these general principles of time translation invariance and linearity is not only cool, it saves you work. As you will see in more complicated examples in a few minutes, it often saves you a LOT of work!

The harmonic oscillator again

Before going on to the harmonic oscillator, let's describe this process in the form of **two simple steps for dealing with systems with time translation invariance and linearity:**[2]

1. Put a trial solution of the form $A\,e^{Ht}$ into your equation of motion. The derivatives become powers of H and you can find the values of $H = h_j$ that work by solving an algebraic equation. This gives you your basis solutions $A\,e^{h_j t}$

2. Write a general solution by making a general linear combination of your basis solutions with arbitrary values of the As —

$$x(t) = \sum_j A_j\,e^{h_j t} \tag{20}$$

   and find the constants A_j by imposing initial conditions, as usual.

Now let's apply these two steps for the mass on the spring. Step 1 is straightforward.

$$m\,\ddot{z}(t) = -K\,z(t) \tag{21}$$

$$z(t) \to A\,e^{Ht} \tag{22}$$

[2]The only time this procedure doesn't work is when the algebraic equation you get has degenerate roots. We will see what to do in this special case later.

$$m\,\frac{d^2}{dt^2}A\,e^{Ht} = -K\,A\,e^{Ht} \quad \text{or} \quad m\,\frac{d^2}{dt^2}A\,e^{Ht} + K\,A\,e^{Ht} = 0 \tag{23}$$

$$m\,H^2\,A\,e^{Ht} + K\,A\,e^{Ht} = 0 \tag{24}$$

$$\left(m\,H^2 + K\right)\,A\,e^{Ht} = 0 \tag{25}$$

$$m\,H^2 + K = 0 = m(H + i\omega)(H - i\omega) \tag{26}$$

$$H = \pm i\omega \quad \text{for} \quad i = \sqrt{-1} \quad \text{and} \quad \omega = \sqrt{\frac{K}{M}} \tag{27}$$

So our two basis solutions are

$$e^{i\omega t} \quad \text{and} \quad e^{-i\omega t} \tag{28}$$

The calculation in step 1 was very simple. The only curious thing here is that we are led to complex numbers. While time translation invariance tells us that there are solutions of the form $A\,e^{Ht}$, it doesn't tell us that H is real, and for oscillations, it isn't.

Now in step 2, we form the general solution by forming a general linear combination of our basis solutions. Thus the most general solution for the harmonic oscillator looks like this:

$$x(t) = c\,e^{i\omega t} + d\,e^{-i\omega t} \tag{29}$$

Now this is a little peculiar. Unlike the situation with the frictional force, in (17), this doesn't look the same as the cosine and sine that we got by solving the differential equation. But in fact, it is the same. The connection with sines and cosines is Euler's formula, one of the more amusing relations in mathematics:

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{30}$$

There are many ways of seeing this. Let's just use the most important formula in physics again — Taylor's expansion

$$e^{i\theta} = 1 + (i\theta) + \frac{1}{2}(i\theta)^2 + \frac{1}{3!}(i\theta)^3 + \frac{1}{4!}(i\theta)^4 + \cdots \tag{31}$$

$$= \left(1 - \frac{1}{2}\theta^2 + \frac{1}{4!}(\theta)^4 + \cdots\right) + i\left(\theta - \frac{1}{3!}\theta^3 + \cdots\right) \tag{32}$$

Because of Euler's formula, you see that the general solution in terms of cosine and sine is completely equivalent to the general solution in terms of complex exponentials.

$$x(t) = a\cos\omega t + b\sin\omega t$$

is equivalent to

$$x(t) = c\,e^{i\omega t} + d\,e^{-i\omega t} \tag{33}$$

$$= c\,(\cos\omega t + i\sin\omega t) + d\,(\cos\omega t - i\sin\omega t)$$

$$a = c + d \qquad b = i(c - d)$$

Notice again how linearity is at work here. Linearity is what guarantees that a linear combination of two possible trajectories is another possible trajectory. This is what allows us to write the most general solution as a combination of the two complex exponential solutions times constants:

$$x(t) = c\,e^{i\omega t} + d\,e^{-i\omega t} \tag{34}$$

It is the fact that the initial conditions appear in this extremely simple way as the coefficients of simple basis solutions that makes all of this work.

If you haven't seen this before, and perhaps even if you have, this probably looks really strange. And you might also be asking yourself, if this is equivalent to the familiar solution in terms of cosine and sine, what is the advantage of using these unfamiliar complex exponential. The right answer, I think, is that once you get used to using complex exponentials, they will simplify your life enormously. We will see this in a few minutes when we discuss damped oscillators, but the message is really more general. We don't have to use complex exponentials. We can do everything using cosines and sines, using a combination of trigonometry and algebra. But with complex exponentials, all we need is algebra!

In fact, Euler's formula is the connection between algebra and trigonometry! You can define the trigonometric functions this way:

$$\cos\theta \equiv \frac{e^{i\theta}+e^{-i\theta}}{2} \tag{35}$$

$$\sin\theta \equiv \frac{e^{i\theta}-e^{-i\theta}}{2i} \tag{36}$$

Now you can derive all trigonometric identities just using algebra, and you never have to do trigonometry again.

Uniform circular motion in the complex plane

One very evocative way to think about these complex solutions is in what is called "the complex plane." Because a complex number has two real components, its real and its imaginary part, we can think of a complex number as a real vector in a two dimensional space in which the real part is the x component of a two dimensional vector and the imaginary part is the y component. This two dimensional space is the complex plane. Euler's formula, (30), tells us that the basis solution $e^{i\omega t}$ has real part $\cos\omega t$ and imaginary part $\sin\omega t$, so its counterpart in the complex plane is the two dimensional vector, $(\cos\omega t, \sin\omega t)$,

$$e^{i\omega t} = \cos\omega t + i\sin\omega t \to (\cos\omega t, \sin\omega t) \tag{37}$$

But this is a unit vector an angle ωt from the x axis. Thus as t increases, $e^{i\omega t}$ executes uniform circular motion in the complex plane. You can see this in the *Mathematica* file **lecture-3-1.nb**.

More generally, a complex number $z = x + iy$ can be written equivalently as a positive number R times a complex exponential $e^{i\theta}$. Note the connection of this with the relation between Cartesian and Polar coordinates in the complex plane.

$$z = x + iy = R\,e^{i\theta} \to (x,y)_{\text{Cartesian}} \Leftrightarrow (R,\theta)_{\text{Polar}} \tag{38}$$

$$R = |z| = \sqrt{x^2+y^2} \tag{39}$$

$$\theta = \arg(z) \tag{40}$$

$$= \begin{cases} \arctan(y/x) \text{ for } x \geq 0\,, \\ \arctan(y/x) + \pi \text{ for } x < 0\,. \end{cases} \tag{41}$$

This connection is used in many practical applications to convert from circular motion to back and forth "reciprocating" motion, or vice versa.[3]

The damped harmonic oscillator

We have now seen what happens when we have friction with a spring or a spring without friction. With a frictional term, the equation of motion for the mass on a spring becomes[4]

$$m\frac{d^2}{dt^2}\,x(t) = -m\Gamma\,\frac{d}{dt}\,x(t) - K\,x(t) \tag{42}$$

$$\frac{d^2}{dt^2}\,x(t) + \Gamma\,\frac{d}{dt}\,x(t) + \omega_0^2\,x(t) = 0 \tag{43}$$

where $\omega_0 = \sqrt{K/m}$. This interpolates between the two systems we have just analyzed. As $\Gamma \to 0$, we recover the harmonic oscillator. As $K \to 0$, we get simple linear friction. But the general form still satisfies the conditions of time translation invariance and linearity. Therefore we still expect the solutions to be an exponentials and we can still use our two steps to construct the general solution.

Step 1 is the usual. Because of time translation invariance and linearity, we can look for exponential basis solutions:

$$z(t) = A\,e^{Ht} \tag{44}$$

$$\frac{d^2}{dt^2}\,A\,e^{Ht} + \Gamma\,\frac{d}{dt}\,A\,e^{Ht} + \omega_0^2\,A\,e^{Ht} = 0 \tag{45}$$

As usual, derivatives with respect to t just bring down factors of H, so we can convert this to an algebraic equation:

$$(H^2 + \Gamma H + \omega_0^2)\,A\,e^{Ht} = 0 \tag{46}$$

Notice that everything has gone exactly the same way here as in our two previous example. Indeed the only difference here is that the algebraic equation, (46), is a little more complicated that the previous examples. We need to use the quadratic formula to find the two solutions:

$$H = -\frac{\Gamma}{2} \pm \sqrt{\frac{\Gamma^2}{4} - \omega_0^2} \equiv H_\pm \tag{47}$$

where

$$\Gamma \geq 0 \Rightarrow \operatorname{Re} H < 0 \Rightarrow \operatorname{Re} H_\pm \leq 0 \tag{48}$$

$z(t)$ oscillates if $\omega_0^2 > \Gamma^2/4$ and it just dies out if $\omega^2 < \Gamma^2/4$ —

$$z(t) \propto e^{H_\pm t} = \overbrace{e^{-\Gamma t/2}}^{\downarrow \text{ with } t} \times \underbrace{\overbrace{e^{\pm t\sqrt{\Gamma^2/4-\omega_0^2}}}^{\downarrow\uparrow \text{ if } \Gamma/2>\omega_0}}_{\bigcirc \text{ if } \Gamma/2<\omega_0} \tag{49}$$

[3]*i.e.* http://www.rpi.edu/dept/chem-eng/Biotech-Environ/PUMPS/reciprocating.html

[4]The little γ that appears in the discussion in Dave's book is just $\Gamma/2$.

Step 2 is also the usual one. The general solution is a linear combination of the simple exponential basis solutions:

$$x(t) = b_+ \, e^{H_+ t} + b_- \, e^{H_- t} \tag{50}$$

where $H_\pm$ are defined in (47). The constants $b_\pm$ contain the information about the initial conditions, just as in the undamped harmonic oscillator. You can see this in the *Mathematica* file **lecture-3-2.nb**.

The *Mathematica* file starts by finding $x(t)$ as a function of ω_0, Γ, and the initial conditions, $x(0)$ and $v(0)$. Then it graphs $x(t)$ with all of these as knobs, starting with $\omega_0 = 0$ — that is pure damping with no spring. The mass is then moving back towards the origin (which is quite arbitrary until we put in the spring). The mass just slows down and stops somewhere (that depends on the initial conditions of course). Now look at what happens if you gradually increase the spring constant, and therefore increase ω_0 with Γ and the initial conditions held fixed. The graph also shows the values of the $H_\pm$ that appear in the exponent of (50) and of the constants $b_\pm$.

Now as we turn on a small ω_0 term by putting in a weak spring, you might think that each of these components would start to oscillate. But that is not what happens. And it cannot possibly work that way because oscillating terms necessarily come in pairs - proportional to $\sin \omega t$ and $\cos \omega t$ or $e^{\pm i\omega t}$ and there cannot be two such pairs with different exponential dependence or we would need more than two initial conditions to determine the trajectory. Instead, the small spring term produces two terms with exponential decay and the decay constants get closer together as ω_0 increases. You can see that this must be so by reminding yourself about where the quadratic formula comes from. The two exponentials, $e^{H_\pm t}$, are determined by (46) as follows:

$$\overbrace{H^2}^{\text{inertial}} + \overbrace{H\,\Gamma}^{\text{frictional}} + \overbrace{\omega_0^2}^{\text{spring}} = 0 = (H - H_+)(H - H_-) \tag{51}$$

$$H_+ + H_- = -\Gamma \qquad\qquad H_+ H_- = \omega_0^2 \tag{52}$$

Evidently, the spring term determines the product of the two Γs while the frictional term determines the sum. The product starts at zero for $\omega_0 = 0$ Not until the product of the Hs get larger than a quarter the square of the sum does real oscillation begin when the Γs develop an imaginary part. At this point, the coefficients $b_\pm$ also become complex. While the mathematics of this is unambiguous, and it is very clear in the *Mathematica* file, I would like to be able describe this more physically.

At any rate, it is clear that the nature of the trajectories described by (50) depends on the relative size of the two parameters, Γ and ω_0. If $\Gamma > 2\omega_0$, the damping is large (this is called "overdamped"). In this case, both H_+ and H_- are real and negative, and the trajectory is a sum of decaying exponentials.

If $\Gamma < 2\omega_0$, the damping is small (this is called "underdamped"). In this case, both H_+ and H_- have a negative real part and an imaginary part (with opposites signs). In this case, the trajectory oscillates (or circles in the complex plane), but also dies out with time exponentially in t. You can see this in the *Mathematica* file **lecture-3-3.nb**.

At the boundary between these two cases — if $\Gamma = 2\omega_0$, the system is "critically damped." In this case, the general solution is

$$x(t) = (A + B\,t)\, e^{-\Gamma\, t/2} \tag{53}$$

Harmonic oscillation for more degrees of freedom

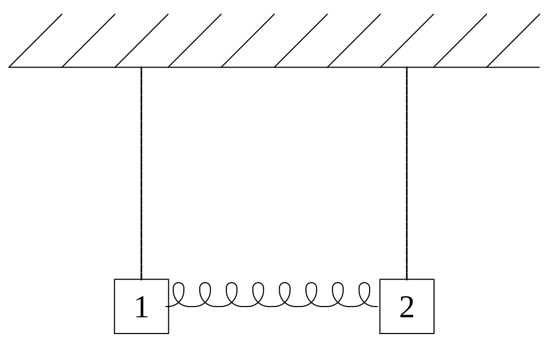
Because our analysis of the harmonic oscillator is very general, relying only on the general principles of linearity and time translation symmetry, the result of (5)

$$z(t) = z(0)\,e^{Ht} \tag{5}$$

applies to any system satisfying these two principles. For example, it is not necessary to restrict ourselves to a single degree of freedom. With more degrees of freedom, $z(t)$ becomes a vector with number of components equal to the number of degrees of freedom, as does $z(0)$ in (5). Thus (5) implies that there are special solutions in which all the components of z move in lockstep, with the same angular frequency. Such a motion is called a "normal mode". The same two steps suffice to solve these more complicated problems, but now there are more basis solutions (because there are more degrees of freedom) and each of the basis solutions describes a motion of ALL the parts of the system. We are not going to explore this in detail in this course. It is the starting point for the study of wave phenomena in Physics 15c. But it is such a useful way of talking about the motion of so many classical systems that I will spend a little time on one important example.

The double pendulum

Here is a very simple example of normal modes that I hope will make the idea clear. Consider the double pendulum, which looks like this:



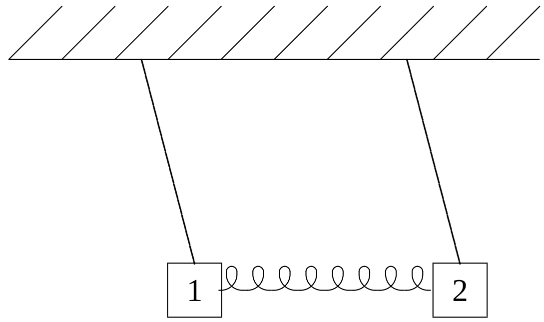
(54)

Two identical pendulums are constrained to move in the plane of the paper and coupled together by a massless spring with spring constant K. In this case the configuration can be labeled by two numbers, x_1 and x_2 the displacements of block 1 and block 2 from equilibrium. Thus this is a system with two degrees of freedom. The vector $q(t)$ that describes the configuration is just

$$q(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \tag{55}$$

Step 1 is now more complicated because we have to find the normal modes. Without the spring, the two pendulums would oscillate independently. For small oscillation, the oscillation of a single pendulum is harmonic with angular frequency $\omega = \sqrt{g/\ell}$. The spring couples these motions together. However, the normal modes are still harmonic. There are two basis solutions for each normal mode.

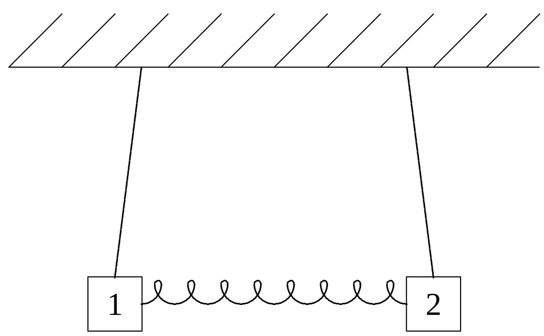
The normal modes look like this:



(56)

There is one normal mode in which the blocks move together. In this mode, the spring in the middle is never stretched from its equilibrium length. The angular frequency of this mode is just the same as the angular frequency of a single pendulum, which (for small oscillations for which the system is linear) is $\omega_1 = \sqrt{g/\ell}$ where ℓ is the distance from the pivot to the mass. The basis solutions that describe this normal mode are

$$z(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\omega_1 t} \quad \text{and} \quad z(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega_1 t} \tag{57}$$



(58)

There is another normal mode in which the two pendulums move in precisely opposite directions. The frequency of this mode is slightly higher than $\sqrt{g/\ell}$, because now the spring contributes to the restoring force that produces the oscillation. In fact, in this case, $\omega_2 = \sqrt{2K + g/\ell}$. The basis solutions that describe this normal mode are

$$z(t) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega_2 t} \quad \text{and} \quad z(t) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-i\omega_2 t} \tag{59}$$

The thing about a normal mode is that the ratios of the displacements of all the parts of the system are fixed throughout the motion. In the double pendulum, in the motion (56), $x_1(t)/x_2(t) = 1$ throughout the motion, while in (58), $x_1(t)/x_2(t) = -1$.

You should have read a bit more about these in Morin's book. In general, it is not easy to find normal modes. In a case like this, you can guess them just from the symmetry of the system. The details of finding them in general involves heavy-duty linear algebra, and this can wait for Physics 15c. What I care about in this course is that you know what these things mean and how to use them if someone tells you what they are.

The point is that step 2 is the same as before. The general solution is always obtained by taking a general linear combination of the basis solutions:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\omega_1 t} + D_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega_1 t} + C_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega_2 t} + D_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-i\omega_2 t} \tag{60}$$

This is short-hand notation for two equations,

$$\begin{aligned} x_1(t) &= C_1\, e^{i\omega_1 t} + D_1\, e^{-i\omega_1 t} + C_2\, e^{i\omega_2 t} + D_2\, e^{-i\omega_2 t} \\ x_2(t) &= C_1\, e^{i\omega_1 t} + D_1\, e^{-i\omega_1 t} - C_2\, e^{i\omega_2 t} - D_2\, e^{-i\omega_2 t} \end{aligned} \tag{61}$$

As usual, we can use Euler's equation to rewrite this in terms of sines and cosines.

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = A_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cos\omega_1 t + B_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \sin\omega_1 t + A_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cos\omega_2 t + B_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \sin\omega_2 t \tag{62}$$

Again, this is short-hand notation for two equations,

$$\begin{aligned} x_1(t) &= A_1\, \cos\omega_1 t + B_1\, \sin\omega_1 t + A_2\, \cos\omega_2 t + B_2\, \sin\omega_2 t \\ x_2(t) &= A_1\, \cos\omega_1 t + B_1\, \sin\omega_1 t - A_2\, \cos\omega_2 t - B_2\, \sin\omega_2 t \end{aligned} \tag{63}$$

lecture 4

Topics:

- Where are we?
- Conservation of Energy
- Work and Energy and the second Law
- Energy in the harmonic oscillator
- More degrees of freedom
- Forced oscillation and resonance
- Harmonic driving forces
- Energy in the driven oscillator
- Breaking the wine glass

Where are we?

We have now seen a number of examples of the use of $F = ma$. You should all be familiar at this point with the techniques of determining the trajectory of a classical system that is picked out by a given set of initial conditions.

We also discussed the beautiful and surprisingly general behavior of the harmonic oscillator, and the particularly simple description of its motion that obtains when we allow our trajectories to involve $i = \sqrt{-1}$.

Now, we will start the process of going beyond $F = ma$ to more general and powerful approaches. This will occupy us for the next two weeks. This week, we will discuss one of the great conservation laws of classical mechanics — conservation of energy.

I also have some organizational remarks. First, a bit of good news — we have now seen all the ways of solving differential equations that we are going to use. There are really only two of them. One is integrating, perhaps after moving things around a little bit to separate variables on two sides of the equation before integrating. The other is hoping that the solution is an exponential, plugging in and checking to see whether it works! Nothing more complicated is going to happen! We will, however, do a little more multivariable calculus today. Please please please stop me and ask questions if you see something you don't understand.

Now to work (literally and figuratively)!

Conservation of energy

It is somewhat unfortunate that "energy conservation" has come to mean two very different things. When we read in a newspaper about energy conservation, the article is usually about using energy carefully and not wasting it. What we mean in classical mechanics when we say energy conservation is something very different. We mean that there is a quantity called "energy" that is unchanged with time for any possible trajectory of the classical system. And in fact this physical meaning of energy conservation seems to be a basic law that survives beyond classical mechanics into the quantum realm. I suppose that there is some connection between these two meanings. If energy were not conserved in the physics sense, if we could simply make new energy whenever we need it, then perhaps we would not have to be so worried about using it sparingly and efficiently. It is the fact that energy is conserved in the physics sense that makes it such an important quantity.

Today and next Tuesday, I will give some examples of the use of conservation of energy. In the weeks that follow, we will address the general issue of conservations laws more systematically and see where they come from at a deeper level. But before we can do that, we will have to learn more about why classical mechanics looks the way it does. This will require that we reformulate classical mechanics in a very beautiful way. And paradoxically, it will involve quantum mechanics as well.

Work and Energy and the second Law

Let's begin by considering a single degree of freedom and a trajectory $x(t)$ satisfying $F = ma$ for some force law. We will often simplify the formulas by not writing the the (t) in $x(t)$ explicitly, but it is important to note that we are considering not a random function, but a trajectory, and if we were to write things out in gory detail, $F = ma$ would be the second order differential equation that $x(t)$ satisfies:

$$F(x(t), \dot{x}(t), t) = m\,\ddot{x}(t) \tag{1}$$

Now we write $F = ma$ as

$$m\,\frac{dv}{dt} = F \tag{2}$$

Multiplying both sides by v gives

$$v\,m\,\frac{dv}{dt} = \frac{d}{dt}\left(\frac{1}{2}m\,v^2\right) = F\,v \equiv P \tag{3}$$

We will define the kinetic energy of a particle of mass m to be $\frac{1}{2}m\,v^2$, thus the left hand side of (3) is the rate of change of kinetic energy. The right hand side is another important quantity called the power, P, supplied by the force to the particle. Thus (3) is the differential form of the work-energy relation. The power supplied by the force acting on a particle is the rate of change of the particle's kinetic energy. This is always true whether there is a conserved energy or not — it follows for any possible trajectory of the system, just from $F = ma$. Integrating (3) with respect to time gives the classic work-energy relation, that the change in kinetic energy equals the work done by the force on the particle:

$$\frac{1}{2}m\,v_2^2 - \frac{1}{2}m\,v_1^2 = \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{1}{2}m\,v^2\right)dt = \int_{t_1}^{t_2} P\,dt = \int_{t_1}^{t_2} v\,F\,dt = \int_{x_1}^{x_2} F\,dx \tag{4}$$

The left hand side of (3) is a total time derivative. If the product $F\,v$ is also a total time derivative, then we can find a quantity that doesn't change with time — its time derivative is zero. To see this, call

$$F\,v = -\frac{d}{dt}U \tag{5}$$

then (3) becomes

$$0 = \frac{d}{dt}\left(\frac{1}{2}m\,v^2\right) - F\,v = \frac{d}{dt}\left(\frac{1}{2}m\,v^2\right) + \frac{d}{dt}U = \frac{d}{dt}\left(\frac{1}{2}m\,v^2 + U\right) \tag{6}$$

Then quantity in parenthesis doesn't change with time and we give it a name, energy,

$$E = \frac{1}{2}m\,v^2 + U \tag{7}$$

and we say that it is "conserved" because it doesn't change with time.

A simple way in which (5) can happen is for F to be a function only of x. Then we can take U in (5) to be

$$U(x) = -\int_{x_0}^{x} F(x')\,dx' \tag{8}$$

Physically, this makes sense because the integral in (8) keeps track of the work done by the force in moving from x_0 to x. If this is negative, then to make the system move in this way, we would have to do work against the force, and would get that energy back by letting the system move back to x_0. Mathematically this works because we can use the chain rule to write

$$\frac{d}{dt}\,U(x) = \frac{dx}{dt}\,\frac{d}{dx}\,U(x) = -\frac{dx}{dt}\,\frac{d}{dx}\,\int_{x_0}^{x} F(x')\,dx' = -v\,F(x) \tag{9}$$

So that $v\,F$ is a total time derivative,

$$v\,F(x) = -\frac{d}{dt}\,U(x) \tag{10}$$

In this case, as I'm sure you know, $U(x)$ is called the potential energy, and the force is just minus the derivative of U,

$$F(x) = -\frac{d}{dx}\,U(x) \tag{11}$$

Note that it is crucial that U depends only on x. Mathematically, you can see that $Fv = -dU/dt$ doesn't work if U is a function of v or t. If U depends on v (like a frictional force), time derivatives of U give terms that depend on the acceleration a - which the force is not allowed to do. If U doesn't depend on v but does depend on t the derivative with respect to t has pieces with no factor of v. You need U to have only x dependence to always get the factor of v from the chain rule in (9).

Energy in the harmonic oscillator

For the harmonic oscillator, the potential energy is

$$U(x) = -\int_{x_0}^{x} F(x')\,dx' = -\int_{x_0}^{x} \left(-Kx'\right)dx' = \frac{1}{2}K\,x^2 - \frac{1}{2}K\,x_0^2 \tag{12}$$

We can take the starting position x_0 to have any fixed value, so we might as well start at the equilibrium position, $x_0 = 0$ so the second term goes away. The second term in (12) doesn't do anything anyway, because it is a constant independent of x so it just drops out when we differentiate to get the force. Thus we can drop the second term and take the conserved energy in the harmonic oscillator to be

$$E = \frac{1}{2}m\,\dot{x}^2 + \frac{1}{2}K\,x^2 \tag{13}$$

It is instructive to see how this works for a general solution of the form

$$x(t) = c\cos(\omega t - \phi) \tag{14}$$

Putting (14) into (13), we get the following for the energy:

$$E = \frac{1}{2}m\,\omega^2\,c^2\sin^2(\omega t - \phi) + \frac{1}{2}K\,c^2\cos^2(\omega t - \phi) \tag{15}$$

Since $\omega = \sqrt{K/m}$, this is

$$E = \frac{1}{2} K\, c^2 \sin^2(\omega t - \phi) + \frac{1}{2} K\, c^2 \cos^2(\omega t - \phi) = \frac{1}{2} K\, c^2 \tag{16}$$

which is a constant, as we expected. But notice the way the mathematics manages to encode the sloshing back and forth of energy between kinetic and potential.

More degrees of freedom

Consider a system with two degrees of freedom with the configuration specified by the two coordinates x_1 and x_2 with masses m_1 and m_2 respectively. Denote the forces on masses 1 and 2 by F_1 and F_2. We will begin by considering the situation in which the forces depend only on the coordinates, not on their derivatives, so that that Newton's second law is

$$m_1 \ddot{x}_1 = F_1(x_1, x_2) \qquad\qquad m_2 \ddot{x}_2 = F_2(x_1, x_2) \tag{17}$$

With one degree of freedom, we saw that there is automatically a conserved energy if the force depends only on position. Here we will see that the situation is more complicated. To see what the issues are, let's compute the time derivative of the kinetic energy,

$$\frac{d}{dt}\left(\frac{1}{2} m_1 \dot{x}_1^2 + \frac{1}{2} m_2 \dot{x}_2^2\right) = m_1 \dot{x}_1 \ddot{x}_1 + m_2 \dot{x}_2 \ddot{x}_2 = \dot{x}_1 F_1(x_1, x_2) + \dot{x}_2 F_2(x_1, x_2) = P \tag{18}$$

P is the power fed into the system by the force. If the right hand side of (18) is the time derivative of something,

$$\dot{x}_1 F_1(x_1, x_2) + \dot{x}_2 F_2(x_1, x_2) \stackrel{?}{=} -\frac{d}{dt} U \tag{19}$$

then we can define an energy that doesn't change with time

$$E = \frac{1}{2} m_1 \dot{x}_1^2 + \frac{1}{2} m_2 \dot{x}_2^2 + U \tag{20}$$

To see when there is a U satisfying (19), we need an important result from multivariable calculus. The change in a function of several variable can be written as a sum of the changes in the variables times the corresponding partial derivatives:

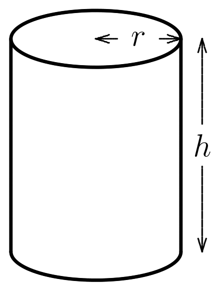
$$df(a, b, c, \cdots) = da\, \frac{\partial f}{\partial a} + db\, \frac{\partial f}{\partial b} + dc\, \frac{\partial f}{\partial c} + \cdots \tag{21}$$

The reason is not too hard to understand. To find the change in f due to a change in a with the other variables held fixed, you would multiply the change in a by the derivative of f with respect to a with the other variables held fixed (which is what the partial derivative means) to get the first term in (21). Similarly, to find the change in f due to a change in b with the other variables held fixed, you would multiply the change in b by the derivative of f with respect to b with the other variables held fixed to get the second term in (21). And so on! Then the total change in f is just the sum of all of these independent changes. When you say it in words rather than writing it down mathematically it makes perfect sense. If you want to know how something changes, you have to add up all the possible sources of change!

We can divide (21) by dt to get the total time derivative,

$$\frac{d}{dt} f(a,b,c,\cdots) = \dot{a}\,\frac{\partial f}{\partial a} + \dot{b}\,\frac{\partial f}{\partial b} + \dot{c}\,\frac{\partial f}{\partial c} + \cdots \tag{22}$$

Here is an example of (22). Consider the volume of a cylinder with height h and radius r.



$$V = \pi\, r^2\, h \tag{23}$$

if both r and h are changing in time

$$\frac{dV}{dt} = \dot{r}\,\frac{\partial V}{\partial r} + \dot{h}\,\frac{\partial V}{\partial h} = \underbrace{\dot{r}\,2\pi\,r\,h}_{\text{change in } r} + \underbrace{\dot{h}\,\pi\,r^2}_{\text{change in } h} \tag{24}$$

Now applying (22) to (19), we see that U had better not depend on derivatives of x_1 and x_2. If it did, (22) would give terms involving $\ddot{x}_1$ and $\ddot{x}_2$, which we don't want. And it should not depend on t, because that would produce a term without an $\dot{x}_1$ or $\dot{x}_2$ because $\dot{t} = 1$. Thus we want

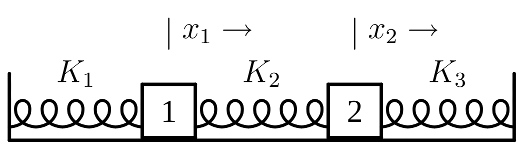
$$\dot{x}_1 F_1(x_1,x_2) + \dot{x}_2 F_2(x_1,x_2) = -\frac{d}{dt} U(x_1,x_2) = -\dot{x}_1 \frac{\partial U}{\partial x_1} - \dot{x}_2 \frac{\partial U}{\partial x_2} \tag{25}$$

which works if

$$F_1(x_1,x_2) = -\frac{\partial U}{\partial x_1} \quad \text{and} \quad F_2(x_1,x_2) = -\frac{\partial U}{\partial x_2} \tag{26}$$

There is a conserved energy if we can get the force by taking partial derivatives of a single function. The function U is the "potential energy."[1] When the forces have the form (26), you can show that the work done by the forces in a motion of the system from one configuration to another is the difference between the initial potential energy and the final potential energy and therefore does not depend on how the system gets from the initial configuration to the final configuration.

Let's do an example of a collection of springs and two masses sliding on a straight frictionless track, as shown in (27).



(27)

We can take the coordinates x_1 and x_2 to be the (signed) distance along the track from the equilibrium positions of the two masses. The forces are[2]

$$F_1(x_1,x_2) = -K_1 x_1 - K_2(x_1 - x_2) \quad \text{and} \quad F_2(x_1,x_2) = -K_2(x_2 - x_1) - K_3 x_2 \tag{28}$$

[1] There could also be contributions to the force that depend on $\dot{x}_1$ and $\dot{x}_2$, but which cancel in the combination $\dot{x}_1 F_1(x_1,x_2) + \dot{x}_2 F_2(x_1,x_2)$. This is the way the magnetic force works.

[2] As always, the direction of "positive" force is the direction in which the coordinate is increasing.

As we might guess from our previous discussion of a single spring, if we choose

$$U(x_1,x_2) = \frac{K_1}{2}x_1^2 + \frac{K_2}{2}(x_1-x_2)^2 + \frac{K_3}{2}x_2^2 \tag{29}$$

then (26) is satisfied. So (29) is the potential energy and the total energy,

$$E = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2 + U(x_1,x_2) \tag{30}$$

is independent of time for any trajectory.

Nothing much changes if the number of degrees of freedom is greater than two, we just have to keep track of more variables. For n degrees of freedom, if we call the coordinates x_j for $j = 1$ to n, the kinetic energy is

$$\sum_{j=1}^{n} \frac{m_j}{2}\dot{x}_j^2 \tag{31}$$

The rate of change of (31) is

$$\frac{d}{dt}\left(\sum_{j=1}^{n} \frac{m_j}{2}\dot{x}_j^2\right) = \sum_{j=1}^{n} m_j\dot{x}_j\ddot{x}_j = \sum_{j=1}^{n} \dot{x}_jF_j(x_1,\cdots,x_n) \tag{32}$$

and if the force on the jth mass is

$$F_j(x_1,\cdots,x_n) = -\frac{\partial}{\partial x_j}U(x_1,\cdots,x_n) \tag{33}$$

then the total energy

$$\left(\sum_{j=1}^{n} \frac{m_j}{2}\dot{x}_j^2\right) + U(x_1,\cdots,x_n) \tag{34}$$

is conserved — that is independent of time for any allowed trajectory,

$$x_1(t)\,,\cdots\,,x_n(t) \tag{35}$$

Forced oscillation and resonance

Suppose we “drive” our damped harmonic oscillator by adding a time dependent force, so the equation of motion becomes

$$\left(\frac{d^2}{dt^2} + \Gamma\,\frac{d}{dt} + \omega_0^2\right)\,x(t) = F(t)/m \tag{36}$$

Let’s begin by discussing the linearity of this equation of motion. Because of the force term, the situation is a bit different from that of an unforced oscillator. Suppose that I have a solution to this equation, $x_1(t)$ and another one $x_2(t)$.

Now if I add them together, I don’t get a solution to the same differential equation

$$
\begin{gathered}
\left(\frac{d^2}{dt^2} + \Gamma \frac{d}{dt} + \omega_0^2\right) x_1(t) = F_1(t)/m \\
\left(\frac{d^2}{dt^2} + \Gamma \frac{d}{dt} + \omega_0^2\right) x_2(t) = F_2(t)/m \\
\Rightarrow \\
\left(\frac{d^2}{dt^2} + \Gamma \frac{d}{dt} + \omega_0^2\right) [a x_1(t) + b x_2(t)] \\
= [a F_1(t) + b F_2(t)]/m
\end{gathered}
\tag{37}
$$

In words, which are not very useful in this case, when we take linear combinations of the solutions, we must also take the same linear combinations of the driving forces, and vice versa.

In particular, this means that can we always add a solution of the homogeneous equation, with no external force.

$$
\begin{gathered}
\left(\frac{d^2}{dt^2} + \Gamma \frac{d}{dt} + \omega_0^2\right) x_1(t) = F_1(t)/m \\
\left(\frac{d^2}{dt^2} + \Gamma \frac{d}{dt} + \omega_0^2\right) x_0(t) = 0 \\
\Rightarrow \\
\left(\frac{d^2}{dt^2} + \Gamma \frac{d}{dt} + \omega_0^2\right) [a x_1(t) + b x_0(t)] = a F_1(t)/m
\end{gathered}
\tag{38}
$$

We can use this form of linearity to simplify the problem.

Harmonic driving forces

Life is much simpler if we look at forces of the form[3]

$$F(t) = F_0\, e^{-i\omega_d t} \tag{39}$$

where ω_d is called the "driving frequency." Why is this an interesting thing to do? This force is exponential — and therefore behaves very simply under time translations.

$$F(t+a) = e^{-i\omega_d a}\, F(t) \tag{40}$$

Thus we can look for solutions that are proportional to $e^{-i\omega_d t}$. There will also be terms in the general solution which are just like those we found for the undriven oscillator. We can always add these solutions because they do not contribute to the driving term. Thus the general solution will be of the form

$$z(t) = \mathcal{A}\, e^{-i\omega_d t} + b_+ e^{-\Gamma_+ t} + b_- e^{-\Gamma_- t} \tag{41}$$

where the

$$e^{-\Gamma_\pm t} \tag{42}$$

[3]We can actually use linearity to write any force as a linear combination of forces of this form using the mathematical technique of Fourier analysis. So if we solve the problem for all values of ω_d, we have actually solved it for all reasonable forces.

with

$$\Gamma_\pm = \frac{\Gamma}{2} \pm \sqrt{\frac{\Gamma^2}{4} - \omega_0^2} \tag{43}$$

are exponential solutions to the unforced oscillator problem. Just as in the case without damping, the coefficients of the "homogeneous" solutions must be set by initial conditions, but that is not true for $\mathcal{A}$. It does not depend on initial conditions! This is what is called a "particular" solution to the differential equation, and the general solution will also involve solutions to the equation without any force, and the coefficients of the solutions will have to be determined by the initial conditions.

One nice thing about the form (41) is that in certain cases, we do not care about the initial conditions. This is because the homogeneous solutions die out exponentially so long as there is any damping at all. If we wait long enough, only the term proportional to $\mathcal{A}$ survives.

Now let's compute $\mathcal{A}$. This is straightforward because we are working with exponentials.

$$\left(\frac{d^2}{dt^2} + \Gamma\,\frac{d}{dt} + \omega_0^2\right) \mathcal{A}\, e^{-i\omega_d t} = F_0\, e^{-i\omega_d t}/m$$

$$(-\omega_d^2 - i\Gamma\omega_d + \omega_0^2)\,\mathcal{A} = \frac{F_0}{m} \tag{44}$$

$$\mathcal{A} = \frac{F_0/m}{\omega_0^2 - i\Gamma\omega_d - \omega_d^2} \tag{45}$$

$$\mathcal{A} = \frac{F_0/m}{\omega_0^2 - \omega_d^2 - i\Gamma\omega_d} \tag{46}$$

$$\mathcal{A} = \frac{F_0/m}{\omega_0^2 - \omega_d^2 - i\Gamma\omega_d} \tag{47}$$

$$\mathcal{A} = \left(\frac{F_0/m}{\omega_0^2 - \omega_d^2 - i\Gamma\omega_d}\right)\left(\frac{\omega_0^2 - \omega_d^2 + i\Gamma\omega_d}{\omega_0^2 - \omega_d^2 + i\Gamma\omega_d}\right) \tag{48}$$

$$\mathcal{A} = \frac{(\omega_0^2 - \omega_d^2 + i\Gamma\omega_d)\,F_0/m}{(\omega_0^2 - \omega_d^2)^2 + \Gamma^2\omega_d^2} \tag{49}$$

Because we used the exponential solution, we got the solution just using algebra. It's $\propto F_0$.

$$\mathcal{A} = \frac{(\omega_0^2 - \omega_d^2 + i\Gamma\omega_d)\,F_0/m}{(\omega_0^2 - \omega_d^2)^2 + \Gamma^2\omega_d^2} = A + iB \tag{50}$$

$$A = \frac{(\omega_0^2 - \omega_d^2)\,F_0/m}{(\omega_0^2 - \omega_d^2)^2 + \Gamma^2\omega_d^2} \tag{51}$$

$$B = \frac{\Gamma\omega_d\,F_0/m}{(\omega_0^2 - \omega_d^2)^2 + \Gamma^2\omega_d^2} \tag{52}$$

Now what do we do with this complex solution. We are not really interested in the complex force we started with. However, we are interested in forces of the form

$$F_0 \cos\omega_d t = \mathrm{Re}(F_0\, e^{-i\omega_d t}) \tag{53}$$

This describes a real harmonic driving force. Now the point is that because of linearity, we can find the solution for the real part of the force by just taking the real part of our complex solution. Taking the real part gives

$$x(t) = \mathrm{Re}\left(\mathcal{A}e^{-i\omega_d t}\right) = A\cos\omega_d t + B\sin\omega_d t \tag{54}$$

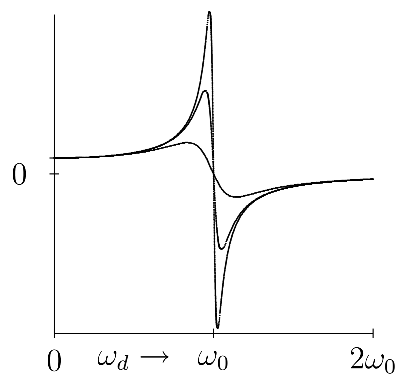
Note the phase relations. The first term is in phase with the force if $\omega_0^2 > \omega_d^2$, that is when the system is driven slowly. In this limit, inertia is irrelevant, and the mass just moves along with the driving force.

If $\omega_0^2 < \omega_d^2$, when the system is driven rapidly, the first term is 180° out of phase with the force. In this limit, inertia dominates.

In between, for $\omega_0^2 = \omega_d^2$, the second term is crucial. It is 90° out of phase (behind) the driving force. This is illustrated in *Mathematica* file **lecture-4-1.nb**.

A and B are called the elastic and absorptive amplitudes for reasons that we will discuss shortly.

Here are graphs of A and B for Γ/ω_d = 0.3, 0.1 and 0.05 For larger values of Γ/ω_d, the resonance hardly shows up at all.



Each of these starts at

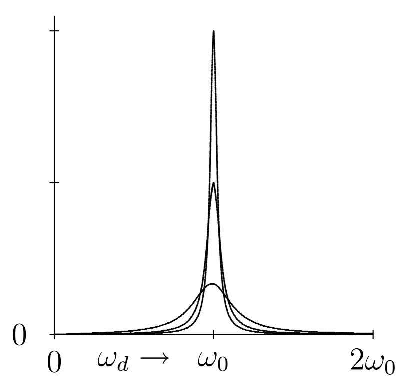
$$\frac{F_0}{m\omega_0^2} \tag{55}$$

for small damping, and looks like

$$\frac{F_0/m}{\omega_0^2 - \omega_d^2} \tag{56}$$

except near resonance.

Now, look at B for the same three values of Γ



Notice that as the Γ/ω_0 decreases, the resonance gets sharper and the effect of the B term is more and more concentrated near the resonance.

Energy in the driven oscillator

To see why A and B in (51) and (52) are called the elastic and absorptive amplitudes, consider the work done by the driving force. The power, $P(t)$, which is the work per unit time done by the driving force, is the force times the velocity of the system on which it acts —

$$P(t) = F(t)\,\dot{x}(t) = F_0 \cos\omega_d t \cdot \frac{\partial}{\partial t}(A\cos\omega_d t + B\sin\omega_d t) \tag{57}$$

Note that the power is a nonlinear function. For example, if F_0 doubles, both $F(t)$ and $\dot{x}(t)$ double, so the power quadruples. Because of this nonlinearity, we cannot use the complex form for $x(t)$ because we could get contributions to the power from both the real and imaginary part, which is not what we want physically. We must use real form for $x(t)$, which is what we have done. Continuing,

$$= P(t) = F_0 \cos\omega_d t \cdot (-\omega_d\, A\sin\omega_d t + \omega_d\, B\cos\omega_d t) \tag{58}$$

$$= -F_0\omega_d A\cos\omega_d t\sin\omega_d t + F_0\omega_d B\cos^2\omega_d t \tag{59}$$

The first term in (59) averages to zero over a complete half-period of oscillation because

$$\cos\omega_d t\sin\omega_d t = \frac{1}{2}\sin 2\omega_d t \tag{60}$$

and

$$\int_{t_0}^{t_0+\pi/\omega_d} \sin 2\omega_d t\, dt = -\frac{1}{2}\cos 2\omega_d t\Big|_{t_0}^{t_0+\pi/\omega_d} = 0 \tag{61}$$

This is why A is called the elastic amplitude. If A dominates, energy that goes in comes back out, like an elastic.

The second term is always positive — it averages to

$$P_{\text{average}} = \frac{1}{2}F_0\omega_d B \tag{62}$$

B is called the absorptive amplitude because it measures how fast energy is absorbed by the system. P_{average}, is maximum on resonance, at $\omega_0 = \omega_d$. Furthermore, if the damping is small, the peak is very very sharp. This is one good way to find resonances.

Breaking the wine glass

The concept of energy is quite useful even in situations where the energy is not obviously conserved. Frictional forces, for example, eat up the energy. In fact, as you probably all know, the energy doesn't go away, but it is converted into heat, and this process cannot be easily reversed, so that we cannot get the energy back into kinetic or potential energy. Thus for example in a damped driven harmonic oscillator started from rest, at first, the work done by the external force goes into increasing the amplitude of the oscillation, and the energy sloshes back and forth between kinetic energy and potential energy. But as the amplitude and therefore the velocity increases, more and more of the power is eaten up by friction. When the system reaches its steady state and the free oscillations have died away, all of the work done by the external force is eaten up by friction (at least when averaged over a complete cycle of oscillation).

I will end with a demonstrations that illustrates quite dramatically the process of resonance in damped, driven oscillation. We will do this in a system with many degrees of freedom — a crystal wine glass. The wine glass is a very complicated system, but its physics is translation invariant and for small oscillations, like almost any other oscillator, it is approximately linear. Thus it has normal modes — motions in which all the parts oscillate with a single frequency. Each of the normal mode frequencies corresponds to a particular way in which the wine glass can ring. The lowest frequency is what dominates if we just ring it.

So what we are going to do is to feed energy into the wine glass at the frequency of its lowest ringing mode. For a good wine glass, the damping is really pretty small — it rings for a long time — so we can feed in a lot of energy and get the lowest mode very excited. So we will actually see how the wine glass is deforming as it rings. In fact, of course, we wouldn't be able to do this without some help, because the frequency is very high, and the motion of the glass would just look like a blur. But what we can do is to use a strobe light tied to the driving frequency so that we illuminate the glass at the same point in the cycle each time, or at least very nearly so. The astonishing thing to me about this demonstration is how much the glass actually deforms before it breaks. If you have never seen this before, I think that you will also be surprised — so here goes.

break the wine glass

lecture 4

Topics:
- Where are we?
- Work and Energy in three dimensions
- Examples of potentials in 3-dimensions
- $1/r^2$ forces and field lines
- The force between spheres
- A particle on a frictionless track

Where are we?

Work and Energy in three dimensions

One of the most useful and interesting applications of the concept of energy is to the motion of a single particle in three dimensions. In this case, the configuration of the system is described by the position vector of particle, $\vec{r}$, and the velocity, the acceleration and the momentum are also vectors, $\vec{v} = \dot{\vec{r}}$, $\vec{a} = \ddot{\vec{r}}$ and $\vec{p} = m\,\vec{v}$.

In a sense, we have already dealt with this because we have talked about system with more than one degree of freedom, and the particle in three dimensions is just a particular example of a system with three degrees of freedom. But it is a very important one, and the fact that the coordinates form a vector in three dimensional space has important consequences for both the physics and the notation.

In three dimensions, Newton's second law is a vector equation,

$$m\,\vec{a} = m\,\dot{\vec{v}} = m\,\ddot{\vec{r}} = \dot{\vec{p}} = \vec{F} \tag{1}$$

We will have much more to say later about what it means to say that something like (1) is a vector equation. For now, we will simply say that a 3-dimensional vector is an object with three independent components, and we will denote the component of $\vec{r}$ by (x, y, z) as you probably did in high school, and the components of an arbitrary vector $\vec{A}$ by (A_x, A_y, A_z). A crucial concept we will use often is the "dot product" of two vectors:

$$\vec{A} \cdot \vec{B} \equiv A_x B_x + A_y B_y + A_z B_z \tag{2}$$

The dot product is important because it has the same value even if we change the components of the individual vectors by rotating to another coordinate system. Here are a few of its most useful properties:

$$\text{The length of a vector } \vec{A} \text{ is } |\vec{A}\,| = \sqrt{\vec{A} \cdot \vec{A}} \tag{3}$$

$$\vec{A} \cdot \vec{B} = |\vec{A}\,|\,|\vec{B}\,| \,\cos\theta_{AB} \;\; \text{where } \theta_{AB} \text{ is the angle between the vectors} \tag{4}$$

$$\frac{1}{|\vec{A}\,|}\,\vec{A} \cdot \vec{B} = |\vec{B}\,| \,\cos\theta_{AB} \;\; \text{is the component of } \vec{B} \text{ in the direction of } \vec{A} \tag{5}$$

Now let's go back to (1). To get the time dependence of the kinetic energy, we take the dot product of both sides with $\vec{v}$,

$$m\,\dot{\vec{v}} \cdot \vec{v} = \frac{d}{dt}\left(\frac{1}{2} m\,\vec{v}^{\,2}\right) = m\,\vec{v} \cdot \vec{a} = \vec{v} \cdot \vec{F} \equiv P \tag{6}$$

This is just the work-energy relation again. The only differences between this and the general situation for many degrees of freedowm is that all three masses are the same, and because of that, the sums can be replaced by dot products. This means that in the power, only the component of the force in the direction of the velocity contributes.

Again, as with one degree of freedom, the time rate of change of the kinetic energy is equal to the power, $P = \vec{v} \cdot \vec{F}$, fed into the system by the force. There is a conserved energy if the power is a total time derivative of a potential:

$$\vec{v} \cdot \vec{F} = -\frac{d}{dt} U \tag{7}$$

But now because the coordinates are the components of a vector, the condition that the forces are the partial derivatives of a single function U can be written as

$$\vec{F}(\vec{r}\,) = -\frac{\partial}{\partial \vec{r}} U(\vec{r}\,) = -\vec{\nabla}\, U(\vec{r}\,) \tag{8}$$

where

$$\vec{\nabla} = \frac{\partial}{\partial \vec{r}} \tag{9}$$

is the "gradient" or "grad" of "div, grad and curl," a vector with components

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \tag{10}$$

We won't make a lot of use of this. You will do more in 15b. But in this case, there is a nice geometrical way of understanding why it is that a force of this kind allows you to define a conserved energy. The work done by the force in going from one point in space to another is independent of the path — it just depends on the starting point and ending point. This is easy to prove, but we won't bother — you will see it later.

The second interesting thing about energy in three dimensions is that we can find velocity dependent forces that still lead to a conserved energy in a very natural way. As long as the force is perpendicular to $\vec{v}$, $\vec{v} \cdot \vec{F}$ is zero, and therefore it is the time derivative of zero, and (7) is satisfied. One important way to do this to use the cross product (which we will discuss in more detail in a few weeks). A force of the form

$$\vec{F} = \vec{v} \times \vec{B} \tag{11}$$

automatically satisfies $\vec{v} \cdot \vec{F} = 0$. This is the way magnetism works.

Examples of potentials in 3-dimensions

It is easy to go from a potential to the force. Equation (8) says that you just differentiate. In fact, this step is so easy that we will often talk about finding the potential by running this backwards — that is finding a $U(x)$ such that (8) gives the force we want.

A familiar example of potential energy is the energy of a massive particle in the earth's gravitational field. Near the surface, the field is roughly constant and vertical, and the force is downward with magnitude $m\,g$. The potential energy is therefore $m\,g\,z$, so that minus the z derivative is the z component of the force, $-m\,g$, and the other components vanish.

$$U(\vec{r}\,) = m\,g\,z \tag{12}$$

$$\vec{F}(\vec{r}) = (0, 0, -m\,g) \tag{13}$$

For another example, consider the potential

$$U(\vec{r}) = \frac{1}{2}\,K\,\vec{r}^{\,2} = \frac{1}{2}\,K\,(x^2 + y^2 + z^2) \tag{14}$$

To get the x, y and z components of the force, we just differentiate and change the sign:

$$\vec{F}(\vec{r}) = (-K\,x, -K\,y, -K\,z) \tag{15}$$

This is a vector that points back towards the origin, so this describes a particle that has a stable equilibrium point at the origin. This is called the three dimensional harmonic oscillator potential.

A more interesting question occurs when you are given a force and are asked to find the potential. It turns out that this is possible if and only if the components of the force satisfy[1]

$$\begin{aligned} \frac{\partial}{\partial y}\,F_z - \frac{\partial}{\partial z}\,F_y &= 0 \\ \frac{\partial}{\partial z}\,F_x - \frac{\partial}{\partial x}\,F_z &= 0 \\ \frac{\partial}{\partial x}\,F_y - \frac{\partial}{\partial y}\,F_x &= 0 \end{aligned} \tag{16}$$

So for example, the force

$$\vec{F}(\vec{r}) = (yz, xz, xy) \tag{17}$$

is associated with a potential because

$$\begin{aligned} \frac{\partial}{\partial y}\,F_z - \frac{\partial}{\partial z}\,F_y &= x - x = 0 \\ \frac{\partial}{\partial z}\,F_x - \frac{\partial}{\partial x}\,F_z &= y - y = 0 \\ \frac{\partial}{\partial x}\,F_y - \frac{\partial}{\partial y}\,F_x &= z - z = 0 \end{aligned} \tag{18}$$

The "only if" part of (16) is easy to understand. If the components of $\vec{F}$ are minus the derivatives of a potential, (16) must be satisfied because the order of differentiations doesn't matter. The "if" part is less obvious. It turns out that if (16) is satisfied, you can find the potential by integrating the force, and it doesn't matter what path you choose to integrate along. This is discussed in Morin's book, and you will see much more of it in Physics 15b. That is enough for now.

[1] As you will see later, this can be written more compactly using the cross product or curl as $\vec{\nabla} \times \vec{F}(\vec{r}) = 0$.

$1/r^2$ **forces and field lines**

Let us start our discussion of gravity by talking about what is so special about $1/r^2$ forces? I will talk both about gravity and also the $1/r^2$ Coulomb force between electrically charged particles. Gravity is really quite different from the Coulomb force. But the mathematics is pretty similar. The potential energy associated with the gravitational force between two massive particles in Newton's theory of gravity is

$$V(\vec{r}_1, \vec{r}_2\,) = -\frac{G\,m_1\,m_2}{|\vec{r}_1 - \vec{r}_2|} \tag{19}$$

where $\vec{r}_1$ and $\vec{r}_2$ are the positions of the particles and m_1 and m_2 are their masses. The force on particle 1 from particle 2 is

$$\vec{F}_{\substack{\text{from } m_2 \\ \text{on } m_1}} = -\frac{\partial}{\partial \vec{r}_1} V(\vec{r}_1, \vec{r}_2\,) = -G\,m_1\,m_2 \frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|^3} \tag{20}$$

Because the potential energy is invariant under space translation because (19) just depends on the vector

$$\vec{r}_{12} = \vec{r}_1 - \vec{r}_2 \tag{21}$$

from $\vec{r}_2$ to $\vec{r}_1$, the satisfies Newton's third law:

$$\vec{F}_{\substack{\text{from } m_1 \\ \text{on } m_2}} = -\frac{\partial}{\partial \vec{r}_2} V(\vec{r}_1, \vec{r}_2\,) = -G\,m_1\,m_2 \frac{\vec{r}_2 - \vec{r}_1}{|\vec{r}_1 - \vec{r}_2|^3} = -\vec{F}_{\substack{\text{from } m_2 \\ \text{on } m_1}} \tag{22}$$

It is also no surprise that the force vector in (20) points from particle 1 to particle 2 because the potential just depends on the distance between the two particles and is therefore invariant under rotations in addition to space translations. So the force has to be proportional to $\vec{r}_{12}$ because there is no other vector in the problem. Then because the force is attractive, pulling particle 1 towards particle 2, the force is a positive factor times $-\vec{r}_{12} = \vec{r}_{21}$.

The potential energy associated with the Coulomb force between two charged particles is (in cgs units)

$$\frac{q_1\,q_2}{|\vec{r}_1 - \vec{r}_2|} \tag{23}$$

where q_1 and q_2 are the charges. Clearly the dependence on the positions is exactly the same. Both potentials lead to a $1/r^2$ force. So we can understand some of the special properties of the gravitational force by thinking about the Coulomb force instead. Before we start, perhaps I should also emphasize the differences. The most obvious difference is that while the electric charge can have either sign, the mass is always positive, so the gravitational force from (23) is always attractive. Less obvious, but even more important, is the effect of the fact that the strength of the gravitational force between two masses is proportional to the masses themselves.[2] This has far-reaching effects, a few of which we will see later.

Now for the similarities and special properties. Newtonian Gravity and the Coulomb force share the crucial property of linearity.[3] The force on a particle due to an arbitrary number of

[2]Note that we are talking here about Newtonian gravity rather than Einstein's General Relativity. Both of these statements must be made more carefully in General Relativity, as we will glimpse after break.

[3]Again this is true of Newtonian gravity but not General Relativity. In General Relativity, nonlinearities are present because the gravitational field carries energy, which in turn produces gravity. We will come back to this later.

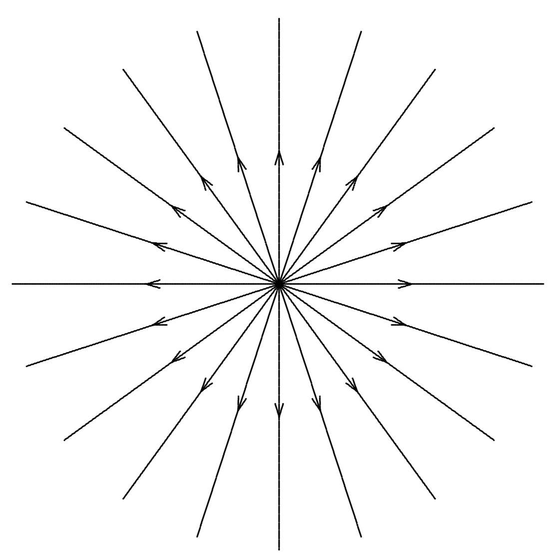
masses or charges is just the sum over the forces from each of the other masses or charges. For gravity

$$\vec{F}_{\text{on } m_1} = \sum_{j \neq 1} \vec{F}_{\substack{\text{from } m_j \\ \text{on } m_1}} = -G\, m_1 \sum_{j \neq 1} m_j \frac{\vec{r}_1 - \vec{r}_j}{|\vec{r}_1 - \vec{r}_j|^3} \tag{24}$$

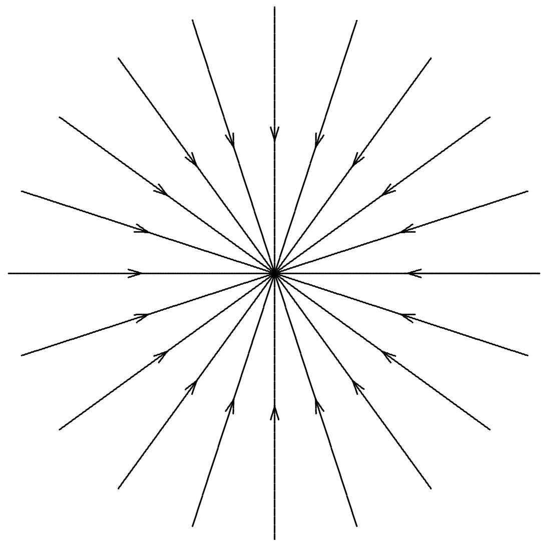
We can extend this in an obvious way to continuous distributions of masses or charges. The sums simply become integrals. In fact, as far as we can tell today, all the masses and charges in the universe are essentially point masses and charges. So our integrals really are approximations to sums. But the individual masses and charges are so small that the so-called continuum approximation, in which we replace a collection of lots of tiny masses or charges with a continuous distribution is often an essentially perfect approximation.

One thing I want to do today is to go over Newton's theorem. It is sufficiently important that it doesn't hurt to look at it several times. In fact, you will see it again in a different and more sophisticated form if you take Physics 15b next semester. The theorem is that the force on a point mass or charge due to a uniform shell of mass or charge is zero inside the shell and outside the shell is just what you would get if all the mass or charge were concentrated at the center. I am not so much interested in a rigorous proof as in making it obvious why the theorem is true.

The key idea that we will use to understand the theorem is something we discussed when we talked about the electromagnetic force. The idea there was "field lines." The electromagnetic force can be thought of in the following way. One charge produces an electromagnetic field that affects other charges, producing the forces (of course, all charges produce these fields, and we don't talk about the effect of the part of the field due to a given charge on the charge that produces it — we will come back to this later in the course). The electric field, in turn, can be thought of as associated with field lines, which begin on positive charges and end on negative charges — they never begin or end in empty space — only on charges. The field lines point in the direction of the electric field, and the density of the field lines is proportional to the strength of the electric field. A single charge sends out field lines symmetrically in all directions. Since the lines can never end if there are no other charged particles around, they continue to spread. On a surface of radius r, the lines are spread out over an area $4\pi\, r^2$. Thus the density of the field lines falls off as $1/r^2$. Thus the electric field and therefore the force falls off as $1/r^2$. This is the basic reason for the Coulomb force.

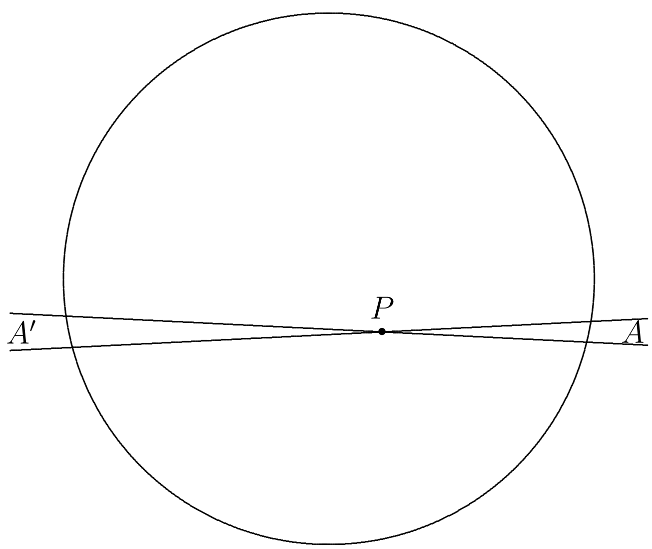


For gravitation, field lines are very similar. Since there is only one kind of charge (which is just mass) and like charges attract, we can think of field lines as coming in from infinity and ending on masses. Like electric field lines, gravitational field lines cannot end where there is no mass. So all the rest of the argument goes through in the same way.



Now what does this mean for Newton's theorem?

First let's show directly that there is no force inside a shell on a point mass at P inside the uniform sphere of mass and look at the the force from P on some infinitesimal area A anywhere on the surface of the sphere. The point is that we can construct an infinitesimal area A' on the other side of the sphere by drawing a straight line from every point on the boundary of A through the point P. The locus of points where these lines intersect the sphere on the other side is the boundary of A'.



(25)

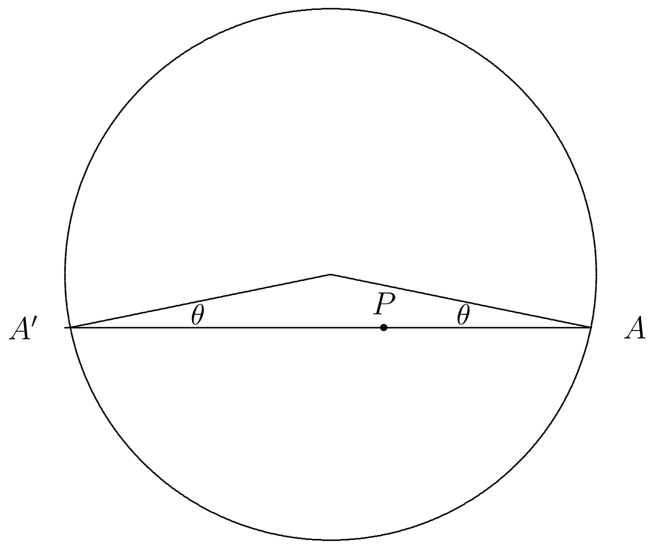
Now we can show that the force on A' is equal and opposite to the force on A.

We can see this geometrically, using the fact that the force goes like $1/r^2$. This is done in problem 5.10 of Morin's book. But field lines make it even more obvious. First look at the force on little areas bounded by the cone at the positions of A and A' but perpendicular to the line between A and A'.

(26)

The number of lines through each area is the same, and since the strength of the field is proportional to the density of the lines and the mass of the little regions is proportional to their area, the number of lines is proportional to the force.

Now the areas of A and A' are larger than this because they are slanted. But both are slanted at exactly the same angle because the triangle formed by the center of sphere and A and A' is isosceles, and the tangents to the sphere are perpendicular to the radii to A and A'.



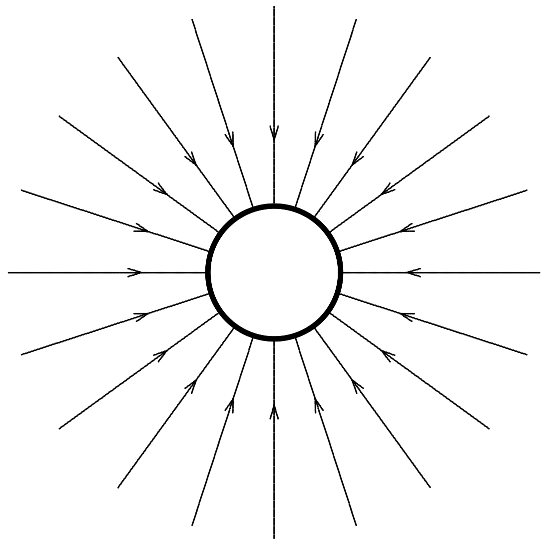
(27)

Thus both areas are larger than the area of the corresponding perpendicular regions by the same factor of $1/\cos\theta$. The the forces are large by the same factor and they are therefore equal and opposite.

But by Newton's third law, the forces on P from the two regions A and A' are equal and opposite.

This shows that for every infinitesimal area on the spherical shell, there is a region on the other side that produces an equal and opposite force at P. This is not quite enough, because we need to be sure that when we integrate over the whole sphere, each of these little areas appears only once. Fortunately, this is easy. We can consider a plane through P and the center of the circle. Then we can find the total force at P by integrating only over the areas on one side of the plane. The areas that cancel the force are all on the other side of the plane. Thus the whole sphere is accounted for, and we have proven the first part of the theorem.

Consider now the field outside a uniform shell of mass. This mass distribution has spherical symmetry. Let us first show that this implies that the force must be radial and equal in all directions. First consider the force at some point $\vec{r}$. This force must be the same if we rotate the mass distribution around the $\vec{r}$ axis, which implies that it must be in the $\vec{r}$ direction because any transverse component would rotate when we rotate the mass distribution. Once we know that the force is radial, it is clear, again because of rotation invariance, that it is the same in all directions, with the magnitude being only a function of $|\vec{r}|$.



Since the force vanishes inside the shell, there is no gravitational field, and thus no field lines inside. All the field lines go from infinity to the shell. But the total mass produces the same number of field lines, whatever the radius of the shell. This must be true because very far away from the shell, it looks just like a point mass, so the number of gravitational field lines coming in from infinity must be the same independent of the radius of the shell. But the field line cannot appear or disappear outside the shell, because there is no mass there. Thus outside the shell, the gravitational field is completely independent of the radius of the shell, and therefore the same as that produced by a point mass at the center. This finishes the theorem.

There are many important consequences of Newton's theorem. One that we will come back to soon is the gravitational force on a point mass m inside a spherically symmetric distribution of masses centered at the origin. If the point mass is at $\vec{r}$ from the center of the distribution, Newton's theorem implies that the force is

$$-\hat{r}\,\frac{G\,m\,M}{r^2} \tag{28}$$

where M is the mass **inside** $\vec{r}$. This will be very important for one of the more mysterious facts about the universe — the existence of dark matter.

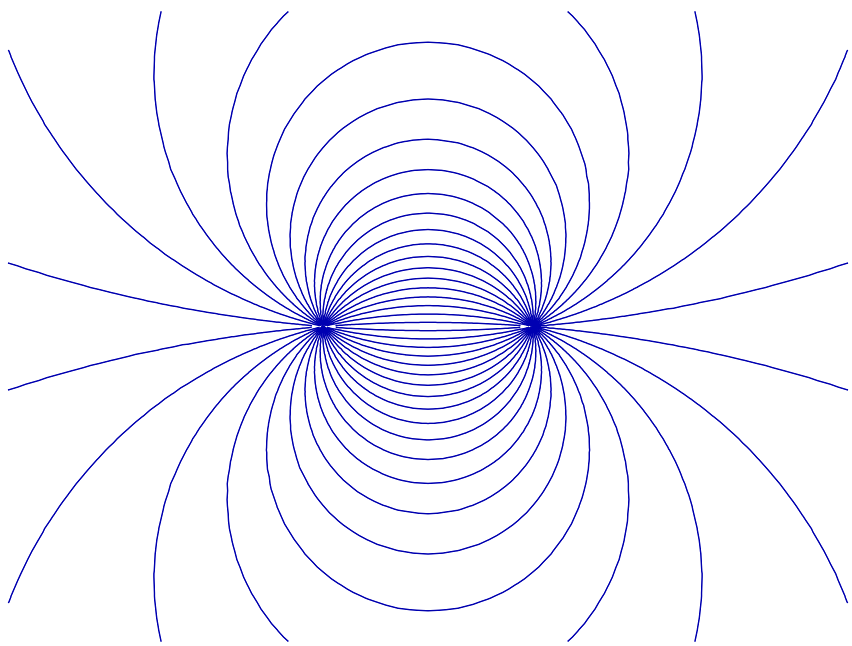
You can think about these arguments involving field lines in two different ways. What is more fundamental - the field lines or the $1/r^2$ force? The straightforward way to interpret this is to say that we know that the force goes like $1/r^2$ and that allows us to show that the field line picture makes sense. Alternatively, we might say that the field lines are the fundamental things, and the field line picture implies (because the field lines spread out over the surface of a sphere which grows like r^2) that the force falls off like $1/r^2$. In fact, what we believe from a more fundamental description of these forces is that the field lines give the more fundamental description. This doesn't matter at all if we stick to three dimensions. But it will be important when we talk later about theories of extra dimensions. In n dimensions, because the area of a sphere of radius r increases like r^{n-1}, the electric and gravitational forces fall like $1/r^{n-1}$!

The force between spheres

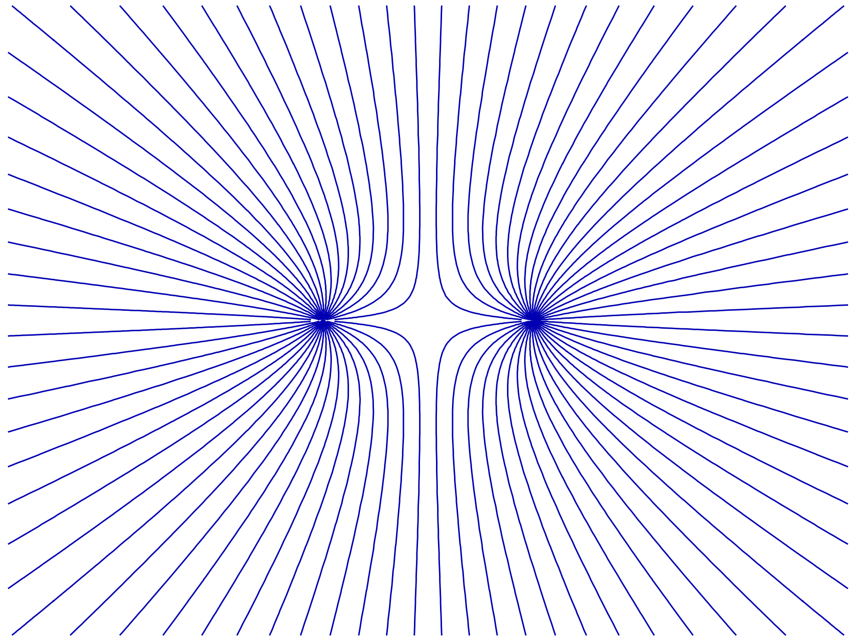
Newton's theorem implies the gravitational force between two spherically symmetric bodies is the same as if all the mass were concentrated at their centers. This follows from linearity and Newton's third law. Consider two spherical bodies, A and B. The force on each of the point masses in B due to A is the same as the force on due to a point at the center of A with the same total mass. But by Newton's third law, this means that the force on A from each of the point masses in B is the same as the force on a point mass of the center of A. But we know that this is the same as a force between a point mass at the center of B and a point mass at the center of A.

This means, among other important things, that spherical bodies exert no torques on one another.

Before we leave (for now) the subject of field lines, I want to show you a couple of plots. One set of field lines that you are probably familiar with is the "dipole" field between positive and negative charges. It looks something like this in terms of field lines:



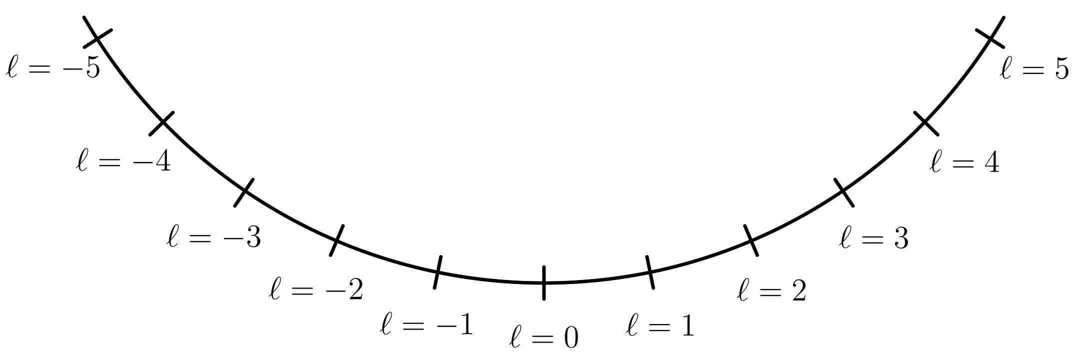
Gravity is different. Because the charges are always the same, the field lines around a pair of equal masses look this (same a electric field lines around a pair of equal negative charges):



The point in the middle is characteristic of a gravitational tidal force. Right in the middle, there is no force. But above or below this point, you are pulled towards the middle, while on either side of the middle, you are pushed away from it. We will come back to this in a couple of months.

A particle on a frictionless track

A system with one degree of freedom need not be one dimensional. You have seen examples of this in the first problem set. Another nice one is a particle on a frictionless track in the earth's gravitational field. Suppose a particle of mass m is constrained to move along a frictionless track that curves around in 3-dimensional space. Because the particle is constrained to be on the track, the configuration of the system is completely specified by the position on the track. It is convenient to measure this in terms of the actual distance of the particle along the track from some reference point on the track, which we will call $\ell(t)$. In terms of $\ell(t)$, we can specify the position of the particle in 3-dimensional space if we have mapped out the x, y and z components of each point on the track as a function of ℓ. In other words, the vector $\vec{r}(\ell)$ describes the shape of the track, and the single variable ℓ describes the position of the particle on the track.



We haven't learned any very simple way to directly write down the equations of motion for such a particle (we will next week). But we can write down the equation for energy, which should be the sum of the kinetic energy and the potential energy due to the earth's gravity.

The speed of the particle is just $\dot{\ell}$, so the kinetic energy should be

$$\frac{1}{2}\, m\, (\dot{\ell})^2 \tag{29}$$

and the potential energy comes from the earth's gravitational field,

$$U = m\, g\, z(\ell) \tag{30}$$

so that the energy is

$$\frac{1}{2}\, m\, (\dot{\ell})^2 + m\, g\, z(\ell) \tag{31}$$

If you think about how difficult it would be to write down the equation of motion and solve it using forces for a complicated shaped track, you will be begin to appreciate the power of conservation of energy.

lecture 6

Topics:
- Where are we now?
- Newton's second law and momentum
- The third law
- Rocket motion
- The Lagrangian and Euler-Lagrange equations
- A better horizontal swing

Where are we?

Last week, we discussed conservation of energy. I will talk about another of the great conservation laws of classical mechanics — conservation of momentum. We will see how it emerges from $F = ma$ and discuss some of its uses. discuss the notion of scattering.

Newton's second law and momentum

Newton's second law for a single particle of mass m can be written as

$$\vec{F} = \frac{d\vec{p}}{dt} \tag{1}$$

where the quantity $\vec{p}$ is the momentum of the particle, and is given in Newtonian mechanics by

$$\vec{p} = m\vec{v}\,. \tag{2}$$

The form (1) actually turns out to be a better and more general way of writing the second law than the familiar $\vec{F} = m\vec{a}$. In this form, (1) [but not (2)] is true even when the speed of the particle approaches the speed of light, where as we will see in a few weeks, many of the common-sense aspects of mechanics begin to break down. In addition, as we will see, (1) often allows us to deal more easily with situations in which objects come apart, or coalesce.

The third law

So if force is always changing momentum according to (1), how is it that momentum is conserved? The answer that you probably learned in high-school is Newton's third law. For every action, there is an equal and opposite reaction. If thing 1 produces a force on thing 2, then thing 2 produces a force with equal magnitude and in the opposite direction on thing 1. If this is correct, then any change of the momentum of something is always compensated by a change in the momentum of the things that are producing the forces on it. The total momentum of any isolated system that has no external forces acting on it is always conserved. For now, you should just accept this. Later in the course (starting next week) we will talk more about why it is true. Here, we will be content to see how it is useful.

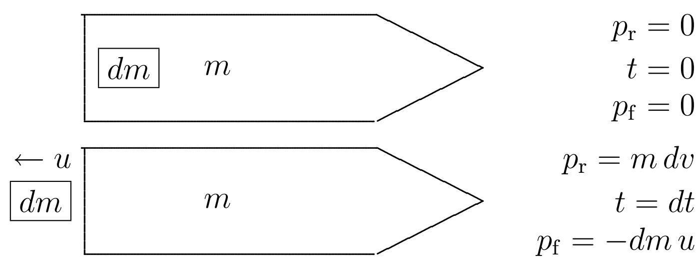
Rocket motion

The most important uses of conservation of momentum all have a couple of things in common. The first is that using momentum conservation allows us to avoid thinking about incredibly complicated details of how the forces work that turn out to be irrelevant in the end. In fact, I personally often find (2) very confusing in situations where the system is changing. Trying to figure out the forces can be very confusing.

It is usually easier and more reliable to make progress instead by comparing the system at two different times and using the fact that the momentum is the same. These two times may be far apart or close together. Sometimes we are interested in the initial condition of a system and the final condition, and we don't much care about what happens in between. But sometimes, we are interested in comparing times that are very close together to use conservation of momentum to analyze the dynamics of the system. In the latter case, we can almost always analyze the problem by looking at the difference between the system at time t and time $t + dt$. A good example of using momentum conservation to simplify the analysis of dynamics is rocket motion. The nozzle of a rocket engine is a very complicated system. There are lots of forces acting on it as the rocket fuel explodes in the nozzle and is forced out at high velocity. If you had to understand in detail the forces acting on the stuff that is ejected from a rocket engine to see how the rocket would move, it would be an impossible job.

But the point is that you never have to mention force at all. We are not interested in the force. There is no good way to measure this force directly. So get rid of it. Conservation of momentum ensures that all you need to know is the velocity, u, of the ejected material, and the rate at which mass is being ejected, dm/dt. You can simply figure out the rate of change in velocity of the rocket, dv/dt, by requiring that momentum be conserved.

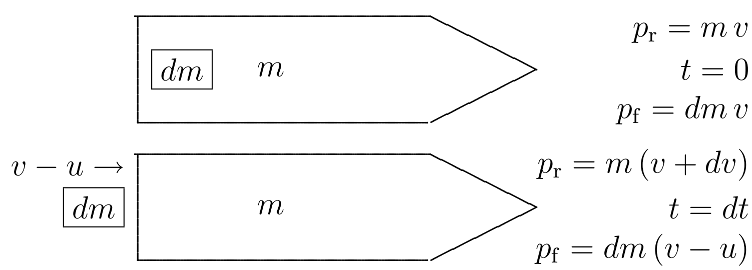
First suppose that the rocket is at rest at time $t = 0$. Then at $t = 0$ the momentum of both the rocket, p_{f}, and the momentum of the fuel, p_{f}, are zero. An infinitesimal time dt later, the momentum of the ejected material is $p_{\mathrm{f}} = -u\,dm$ and this must be compensated by the change in momentum of the rocket in the opposite direction, $p_{\mathrm{r}} = m\,dv$.



Thus

$$m\,\frac{dv}{dt} = \frac{dm}{dt}\,u \tag{3}$$

This is also true if the rocket is moving with velocity v.



The changes in momentum are the same, though the total momentum has changed.

If you are really more comfortable talking about forces, you can use conservation of momentum to calculate the total force on the rocket. then you know that the rate of change of momentum of the ejected material is $u\,dm/dt$. Since this momentum is being produced, the rocket must be pushing on the stuff with a force

$$F = \frac{dp}{dt} = u\,\frac{dm}{dt} \tag{4}$$

$F = u\,dm/dt$. And therefore, according to Newton's third law, the stuff is pushing back on the rocket with a total resultant force of this magnitude.

Notice again that the rate of change of the rocket's momentum is independent of how the rocket is moving (remember that u is the velocity of the ejected fuel **with respect to the rocket**). This means that the force on the rocket looks the same in any inertial frame. This is an example of an important principle in Newtonian mechanics - Galilean Relativity. Galilean Relativity is the statement that Newton's laws are valid in any inertial frame. Forces cannot change in going from one inertial frame to another, because the force is proportional to the acceleration, and the acceleration depends only on the rate of **change** of velocity. Any fixed velocity of the the inertial frame just cancels out when we compute the acceleration.

Thus people in the rocket cannot tell how fast they are going without looking outside. Velocity is only defined in a particular inertial frame. But acceleration is something that the people in the rocket can feel because they can feel the force. The floor of the rocket is pushing against them to accelerate them along with the ship and they feel that force in their bones. But it is the same in any inertial frame. Thus the passengers can tell how they are accelerating, but they cannot tell how fast they are going without looking out the window (or integrating the acceleration from the beginning of the trip).

The Lagrangian and the Euler-Lagrange equations

In addition to discussing momentum conservation this week, we are going to begin our study of a remarkable reformulation of classical mechanics, based on the Lagrangian and the Euler-Lagrange equations. This will replace $F = ma$ as the equation of motion of a classical system with something that is actually easier to use in most of the situations that really matter. This week, we are just going to write down the answer. Next week, we will talk bit about why it works, and why it is so interesting. Our reformulation is not just a great labor saving device, it is actually telling us something very deep and important about the way the world work. This will move us closer to understanding $F = ma$. But already this week, you should begin to see that the Lagrangian of a classical system and the Euler-Lagrange equations that we derive from it are **better** than $F = ma$ in many ways.

So without further ado, let me write down the answer for a single degree of freedom. We begin by constructing the **Lagrangian**, $\mathcal{L}$.

$$\mathcal{L}(x, \dot{x}, t) = T(x, \dot{x}, t) - U(x, \dot{x}, t) \tag{5}$$

Where T and U are the kinetic and potential energies, as functions of the variable x that describes the configuration of the system, and its time derivative, $\dot{x}$. This time I have included the possibility that these functions also depend explicitly on t. By explicit time dependence, I mean dependence beyond that coming from the time dependence of $x(t)$. We will come back to this a little later when we discuss energy. In terms of $\mathcal{L}$, the equation of motion of the system is

$$\frac{\partial}{\partial x}\mathcal{L}(x,\dot{x},t) - \frac{d}{dt}\frac{\partial}{\partial \dot{x}}\mathcal{L}(x,\dot{x},t) = 0 \tag{6}$$

Solving the differential equation (6) and imposing the initial conditions then gives us the classical trajectory $x(t)$.

I will illustrate this today with two examples. The first is a problem you already how to deal with - a mass m moving in force derived from a potential $U(x)$. In this case the kinetic energy depends only on the velocity

$$T(\dot{x}) = \frac{1}{2}m\dot{x}^2 \tag{7}$$

and so the Lagrangian depends only on x and $\dot{x}$:

$$\mathcal{L}(x,\dot{x}) = \frac{1}{2}m\dot{x}^2 - U(x) \tag{8}$$

So

$$\frac{\partial}{\partial x}\mathcal{L}(x,\dot{x}) = -\frac{d}{dx}U(x) \tag{9}$$

$$\frac{\partial}{\partial \dot{x}}\mathcal{L}(x,\dot{x}) = m\dot{x} \tag{10}$$

and the Euler-Lagrange equation is

$$\frac{\partial}{\partial x}\mathcal{L}(x,\dot{x}) - \frac{d}{dt}\frac{\partial}{\partial \dot{x}}\mathcal{L}(x,\dot{x}) = -\frac{d}{dx}U(x) - \frac{d}{dt}m\dot{x} = -\frac{d}{dx}U(x) - m\ddot{x} = 0 \tag{11}$$

or

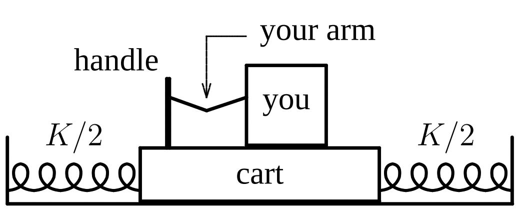
$$-\frac{d}{dx}U(x) = m\ddot{x} \tag{12}$$

or, because minus the derivative of the potential is the force,

$$F = ma \tag{13}$$

A better horizontal swing

OK, so we showed that our funny looking Euler-Lagrange equation is equivalent to $F = ma$. So what? That and about ten bucks will get you dessert at Finales. But now let us do a more interesting problem related to one of the problems you just did.



The system shown above is a cartoon of a different kind of horizontal swing set from the one you studied in AS2. A cart with mass μ slides without friction in the x direction, attached to springs as shown. You with mass m sit on the cart. You and the cart are at rest in equilibrium until $t = 0$. For $t > 0$, by pushing and pulling on the handle with your arm (or arms if you prefer), you slide back and forth on the cart, so that your horizontal position with respect to the handle is

$$\ell \cos \omega_d t \tag{14}$$

where

$$\omega_d = \sqrt{\frac{K}{\mu + m}} \tag{15}$$

for some constant ℓ. Call the horizontal displacement of the cart from equilibrium x. Let's find the equation of motion for this system.

One of the differences between this system and the one you studied on the problem set is that there is only one degree of freedom here — x. Your position with respect to the cart is completely fixed by the statement of the problem, so once we have determined $x(t)$, we know everything.

Another difference is that the resonant frequency depends on the total mass $\mu + m$, rather than just the mass of the cart.

Yet another difference is that we didn't mention the force on you or the handle. We don't need to know it to do the problem. Since we don't know the force, we can't use $F = ma$ directly (though you with sufficient ingenuity, you can find a way to do it). The the Lagrangian technique works just fine.

The potential energy of the system is due to the spring, and we can take it to be

$$\frac{1}{2} K x^2 \tag{16}$$

The kinetic energy of the system gets two contributions, one from the motion of the cart and one from your motion. The kinetic energy of the cart is just

$$\frac{1}{2} \mu \dot{x}^2 \tag{17}$$

Your motion is a little more complicated. Up to a constant, your horizontal position for $t > 0$ is

$$x + \ell \cos \omega_d t \tag{18}$$

and thus your velocity is

$$\dot{x} - \ell \omega_d \sin \omega_d t \tag{19}$$

and your kinetic energy is

$$\frac{1}{2} m (\dot{x} - \ell \omega_d \sin \omega_d t)^2 \tag{20}$$

The total kinetic energy is then

$$\frac{1}{2} \mu \dot{x}^2 + \frac{1}{2} m (\dot{x} - \ell \omega_d \sin \omega_d t)^2 \tag{21}$$

and the Lagrangian is

$$\mathcal{L}(x, \dot{x}, t) = \frac{1}{2} \mu \dot{x}^2 + \frac{1}{2} m (\dot{x} - \ell \omega_d \sin \omega_d t)^2 - \frac{1}{2} K x^2 \tag{22}$$

Now we can write down the Euler-Lagrange equation.

$$\frac{\partial}{\partial \dot{x}}\mathcal{L} = \mu\dot{x} + m(\dot{x} - \ell\omega_d \sin\omega_d t) \tag{23}$$

$$\frac{\partial}{\partial x}\mathcal{L} = -Kx \tag{24}$$

Thus the Euler-Lagrange equation is

$$\frac{d}{dt}(\mu\dot{x} + m\dot{x} - m\ell\omega_d \sin\omega_d t) = -Kx \tag{25}$$

or

$$(\mu + m)\ddot{x} = -Kx + m\ell\omega_d^2 \cos\omega_d t \tag{26}$$

So at the end of the day, we see that this is a forced oscillation problem like the problem on the problem set, except that it involve the sum of the masses. This means that indeed we chose the right driving frequency to get resonance in (15). We can now put in the same form for a trial particular solution that you explored in the problem set:

$$x(t) = A\,t\sin\omega_d t \tag{27}$$

which satisfies

$$\ddot{x}(t) = (-A\omega_d^2\,t\sin\omega_d t + 2A\omega_d\cos\omega_d t) \tag{28}$$

So

$$(\mu + m)(-A\omega_d^2\,t\sin\omega_d t + 2A\omega_d\cos\omega_d t) = -K\,A\,t\sin\omega_d t + m\ell\omega_d^2\cos\omega_d t \tag{29}$$

and

$$A = \frac{m\ell\omega_d}{2(\mu + m)} \tag{30}$$

This system is animated in the *Mathematica* notebook **swing-horizontal2.nb** bundled with your lecture notes.

lecture 7

Topics:
- Where are we now?
- Scattering and kinematics
- Elastic collisions
- Inelastic collisions
- Generalized Force and Momentum
- Example - bead on a expanding ring
- Example - bead on a rotating rod
- More degrees of freedom

Where are we now?

Momentum conservation and Lagrangians are two fantastic labor-saving devices. Next time we will see how they are related. This time we will just do a lot of examples.

Scattering and kinematics

The idea of scattering is very important in many subfields of science (and pseudo-sciences like economics for that matter). The idea is that one does not always have to follow the trajectories of all the particles in a process in detail to learn something about the process. Often, important information can be obtained by just looking at the initial state and the final state, and not asking about the details of what happens in between. Here is an example. Suppose that we have two particles in three dimensional space interacting through a potential that depends only on the difference between the position vectors of the two particle. The energy is then

$$\frac{m_1}{2}\,\dot{\vec{r}}_1^{\,2}+\frac{m_2}{2}\,\dot{\vec{r}}_2^{\,2}+V(\vec{r}_1-\vec{r}_2) \tag{1}$$

We will discuss in more detail later in the course why such a system conserves energy and momentum. For now, just note that this potential energy leads to forces that are consistent with Newton's third law. The force on particle 1 is

$$\vec{F}_1=\left(-\frac{\partial}{\partial x_1}V(\vec{r}_1-\vec{r}_2),-\frac{\partial}{\partial y_1}V(\vec{r}_1-\vec{r}_2),-\frac{\partial}{\partial z_1}V(\vec{r}_1-\vec{r}_2)\right) \tag{2}$$

The force on particle 2 is

$$\vec{F}_2=\left(-\frac{\partial}{\partial x_2}V(\vec{r}_1-\vec{r}_2),-\frac{\partial}{\partial y_2}V(\vec{r}_1-\vec{r}_2),-\frac{\partial}{\partial z_2}V(\vec{r}_1-\vec{r}_2)\right) \tag{3}$$

The chain rule implies that $\vec{F}_1=-\vec{F}_2$, so that Newton's third law is satisfied, and momentum is conserved.

Let's also assume that V is "short-range" which means that

$$V(r)=0\ \text{ for large } |r| \tag{4}$$

or at least that $V(r)$ goes to zero very rapidly as $|r| \to \infty$. Even if we do not know what V looks like in detail, we can say interesting things about this system because energy and momentum are conserved. Now suppose that we consider a process in which particle 2 is at rest at the origin and particle 1 approaches from far away. Initially, the potential is irrelevant because the particles are far apart. Particle 1 has velocity $\vec{v}_{1i}$. After the interaction, the two particles will typically be far apart again, so again the potential will be irrelevant. The importance of the fact that the potential energy is irrelevant is that it means that the total kinetic energy long before the collision is the same as the total kinetic energy long after the collision. At some intermediate time, when the particles were close together, some kinetic energy was converted to potential, but it all comes back as the particles separate. Thus, as long as one only asks about the initial and final states, **kinetic energy** is conserved. And of course, because Newton's third law is satisfied, momentum is conserved at all times, and in particular for the initial and final states. This is a scattering process. We say that the two particle have scattered from one another.

Now the point is that kinetic energy and momentum conservation put very strong constraints on this process. Suppose that particle 1 has velocity $\vec{v}_{1f}$ after the interaction and particle 2 has velocity $\vec{v}_{2f}$.

Since the initial momentum and energy of particle 2 are zero, we can write:

$$m_1\,\vec{v}_{1i} = m_1\,\vec{v}_{1f} + m_2\,\vec{v}_{2f} \qquad \frac{m_1}{2}\,\vec{v}_{1i}^{\,2} = \frac{m_1}{2}\,\vec{v}_{1f}^{\,2} + \frac{m_2}{2}\,\vec{v}_{2f}^{\,2} \tag{5}$$

These constraints are very powerful. If you don't know the mass of particle 2, for example, you can calculate it from the three velocities. Scattering gives you information about the particles involved.

There is an interesting linguistic distinction that is made in calculations like this. We talk about the features that follow from very general principles like conservation of kinetic energy and momentum as the "kinematics" of the process. This is to be distinguished from the "dynamics" of the process, which is everything else — in particular the details of the force law. At this point, the distinction probably seems a little arbitrary, but in the next couple of weeks, as we begin to see how general these conservations laws really are, this distinction will be more and more important.

Another place where the idea of scattering is crucial is in my own field of particle physics. I study particles that are very small. We can detect them, we can see their tracks, measure their velocities, and energies and momenta, just as we would with a larger object. But they are so small, that we cannot follow what happens when two of them collide in detail. We simply cannot measure the forces involved in the tiny fraction of a second during which colliding subatomic particles are in "contact" with one another.[1] What we do is scattering experiments, in which we measure the initial energies and momenta of the particles before the collision, and then again after the collision. Here, conservation of energy and momentum are really useful, because they put very strong limits on what can happen. We will discuss this a bit now, and then in much more detail in a few weeks when we discuss energy and momentum in relativity.

[1] In fact, the whole notion of these particles being in contact with one another is rather problematic. It is not clear what it means.

Elastic collisions

A familiar example of the use of kinematics is in elastic collisions of rigid bodies, like billiard balls. The force in this case is certainly short-range, because it is only non-zero when the balls are actually touching. “Elastic” is just a code word meaning that kinetic energy is conserved.

A famous and beautiful result that follows simply from conservation of kinetic energy and momentum in an elastic collision is that if a moving particle with velocity $\vec{v}$ collides elastically with a particle at rest with the same mass, the dot product of the velocities, $\vec{v}_1$ and $\vec{v}_2$, of the two particles in the final state vanish, $\vec{v}_1 \cdot \vec{v}_2 = 0$. Thus either one of the two velocities vanishes, or else the two velocity vectors are perpendicular to one another. This is very neat result, and it is easy to prove using conservation of kinetic energy and momentum. Conservation of kinetic energy gives

$$\frac{1}{2}m\,\vec{v}^{\,2} = \frac{1}{2}m\,\vec{v}_1^{\,2} + \frac{1}{2}m\,\vec{v}_2^{\,2} \tag{6}$$

Conservation of momentum gives

$$m\,\vec{v} = m\,\vec{v}_1 + m\,\vec{v}_2 \tag{7}$$

Eliminating the ms and taking the dot product of each side of (7) with itself gives

$$\vec{v}^{\,2} = (\vec{v}_1 + \vec{v}_2) \cdot (\vec{v}_1 + \vec{v}_2) = \vec{v}_1^{\,2} + \vec{v}_2^{\,2} + 2\,\vec{v}_1 \cdot \vec{v}_2 \tag{8}$$

Comparing this with (6) gives the desired result. No forces were ever mentioned. In fact, it does not even matter here which of the two particles in the final state was initially at rest and which was moving, because they have the same mass.

There is another way of thinking about this result that is rather neat. Suppose that we look at the process in the zero momentum frame. We can get to this frame by moving all of our measuring apparatus with velocity $\vec{v}/2$. In this new frame, particle that was moving now has velocity

$$\vec{v}_{CM} = \vec{v}/2 \tag{9}$$

and the particle that was initially at rest has the opposite velocity,

$$-\vec{v}_{CM} = -\vec{v}/2 \tag{10}$$

so the sum of the momentum vectors is zero. After the collision, again the particles must be moving with equal and opposite velocities, and energy conservation requires that the speeds be the same as the initial speeds. That is all the particles can do is to change direction without changing speed.

$$\vec{v}_{1CM} = -\vec{v}_{2CM} \quad \text{with} \quad v_{CM}^2 = v_{1CM}^2 \tag{11}$$

Now we can go back to the original frame by just adding back the $\vec{v}_{CM}$ from the motion of the frame. Then you can see that

$$\vec{v}_1 = \vec{v}_{1CM} + \vec{v}_{CM} \qquad \vec{v}_2 = \vec{v}_{2CM} + \vec{v}_{CM} = -\vec{v}_{1CM} + \vec{v}_{CM} \tag{12}$$

Then taking the dot product gives

$$\vec{v}_1 \cdot \vec{v}_2 = \left(\vec{v}_{1CM} + \vec{v}_{CM}\right) \cdot \left(-\vec{v}_{1CM} + \vec{v}_{CM}\right) = -v_{1CM}^2 + v_{1CM}^2 = 0 \tag{13}$$

Inelastic collisions

One of the hard parts of doing physics is figuring out what principles to use in a particular problem. At some level, for example, we believe that energy and momentum are always conserved. Sometimes, as in the example we just discussed, kinetic energy and momentum are obviously conserved. But for example, when I take an egg and drop in on the floor, it is certainly not obvious. Before it hits the floor, the egg has kinetic energy and momentum. After it hits the floor, it is just a scrambled mess, sitting still, with no kinetic energy or momentum. What happened? Clearly the problem with the eggs momentum is that the egg is not an isolated system. The floor pushed up on the egg's shell (and broke it), to change the eggs momentum. The problem with the egg's energy is that energy can get transformed from one form to another. In the process of breaking, the egg's parts actually heated up slightly, so slightly that we don't notice it, but enough so that energy is conserved.

Frequently, we are interested in collisions in which momentum is conserved, but kinetic energy is not. These are called "inelastic" collisions. The classic inelastic collision is two lumps of clay hitting and sticking. If the two lumps have masses m_1 and m_2 with velocities v_1 and v_2 respectively, the final momentum is

$$m_1\vec{v}_1 + m_2\vec{v}_2 \tag{14}$$

We can now compute the final velocity of the system by dividing by the total mass,

$$\vec{v}_{\text{final}} = \frac{m_1\vec{v}_1 + m_2\vec{v}_2}{m_1 + m_2} \tag{15}$$

There are two keys to obtaining (15). Notice that we have not used conservation of energy. What we used instead is the physical picture of the event. The key word is "sticking." This means that the two lumps are moving with the same velocity after the collision. This allows us to use conservation of momentum alone to find the final velocity. There is also an unspoken assumption behind (15)— the common sense statement that mass is conserved, like energy and momentum — that the mass of the system of two lumps stuck together is just the sum of the masses of the two lumps. This is very reasonable, and it is consistent with what each of knows in our bones about the world. However, it is wrong. We will see in a few weeks that when particles collides at speeds close to the speed of light, conservation of mass may be completely wrong. Nevertheless, it is a very good approximation as long as none of the particles are moving at close to the speed of light, and because the speed of light is so enormous compared to what are used to, this is usually not much of a restriction.

Since we haven't used conservation of energy, it is of some interest to calculate how much energy is lost in the collision. This is

$$\frac{m_1}{2}\vec{v}_1^2 + \frac{m_2}{2}\vec{v}_2^2 - \frac{m_1 + m_2}{2}\left(\frac{m_1\vec{v}_1 + m_2\vec{v}_2}{m_1 + m_2}\right)^2 = \frac{\mu}{2}\left(\vec{v}_1 - \vec{v}_2\right)^2 \tag{16}$$

where μ is an important quantity called the "reduced mass." We will discuss more of this later. Note that the energy loss never vanishes in this process unless the lumps were moving together in the first place. Some energy is always lost whenever there is "sticking."

Generalized Force and Momentum

For a particle of mass m moving in a potential $V(x)$, the Euler-Lagrange equation of motion can be written as

$$\frac{dp}{dt} = \frac{d}{dt}(m\dot{x}) = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{x}} = \frac{\partial\mathcal{L}}{\partial x} = -V'(x) = F(x) \tag{17}$$

— the rate of change of the momentum is equal to the force. In the more general situation, this suggests that we might regard the Euler-Lagrange equations,

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_j} = \frac{\partial\mathcal{L}}{\partial q_j} \tag{18}$$

as a generalization of this — we call

$$\frac{\partial\mathcal{L}}{\partial\dot{q}_j} \tag{19}$$

the "generalized momentum" corresponding to the coordinate q_j and

$$\frac{\partial\mathcal{L}}{\partial q_j} \tag{20}$$

the "generalized force" corresponding to the coordinate q_j. Then the Lagrange equation says that the rate of change of the generalized momentum equals the corresponding generalized force.

A particularly interesting case occurs when the Lagrangian does not depend at all on some coordinate q_j. In that case, (18) implies that the generalized momentum corresponding to q_j is constant. This statement becomes even more interesting when you realize that we have great freedom to choose the coordinates any way we want to. Thus if there is **any** coordinate system in which the Lagrangian does not depend on some coordinate, then there is a conservation law — the corresponding generalized momentum is conserved.

Example - bead on an expanding ring

Let's warm up by doing a couple of examples of Euler-Lagrange equations for systems with a single degree of freedom. I'll begin with a system that is mathematically very simple, but that I don't actually know how to build. Suppose we have a small bead with mass m that slides without friction on circular ring centered in the $x-y$ plane, but whose radius grows as a function of time as

$$r(t) = r_0 + v_{rr}\,t \tag{21}$$

It will be most useful to analyze this in terms the polar angle θ in terms of which the x and y coordinates of the bead are

$$x = r(t)\cos\theta \qquad y = r(t)\sin\theta \tag{22}$$

We can find the kinetic energies by differentiating (22) to get the components of the velocity,

$$v_x = v_{rr}\cos\theta - \dot{\theta}\sin\theta \qquad v_y = v_{rr}\sin\theta + \dot{\theta}\cos\theta \tag{23}$$

We can also write the kinetic energy of the bead directly in terms of θ using the fact that the radial motion of the bead due to the expansion of the ring and tangential motion associated with changing θ are instantaneously perpendicular. Either way se see that $\vec{v}^2$ can be written as

$$\vec{v}^2 = \dot{x}^2 + \dot{y}^2 = v_{rr}^2 + r(t)^2\,\dot{\theta}^2 \tag{24}$$

Remember here that $r(t)$ is not a dynamical variable — the time dependence of r is imposed on the system from the beginning. There is no potential energy because the system is in a horizontal plane, so the Lagrangian is

$$\mathcal{L}(\theta,\dot{\theta},t) = \frac{m}{2}\left(v_{rr}^2 + r(t)^2\,\dot{\theta}^2\right) \tag{25}$$

The Euler-Lagrange equation here is particularly simple because the Lagrangian does not depend at all on θ. Thus

$$0 = \frac{\partial}{\partial\theta}\mathcal{L}\Big(\theta,\dot{\theta},t\Big) - \frac{d}{dt}\frac{\partial}{\partial\dot{\theta}}\mathcal{L}\Big(\theta,\dot{\theta},t\Big) = -\frac{d}{dt}\frac{\partial}{\partial\dot{\theta}}\mathcal{L}\Big(\theta,\dot{\theta},t\Big) = -\frac{d}{dt}\Big(m\,r(t)^2\,\dot{\theta}\Big) \tag{26}$$

What is nice about (26) is that the solution (or more properly the "first integral") is really simple.

$$m\,r(t)^2\,\dot{\theta} = \text{a constant} \tag{27}$$

The relation (27) is an example of conservation of a generalized momentum. The Lagrangian does not depend on θ, only on $\dot{\theta}$, so the generalized momentum

$$\frac{\partial\mathcal{L}}{\partial\dot{\theta}} \tag{28}$$

is constant. In fact, as we will discuss in more detail later, this is the angular momentum, which is conserved because there is no torque on the mass about the origin.

To find the constant, we need an initial condition. For example if the angular velocity at $t = 0$ is $\dot{\theta}(0) = \omega_0$, then we can write

$$m\,r(t)^2\,\dot{\theta} = m\,r_0^2\,\omega_0 \tag{29}$$

thus

$$\dot{\theta} = \frac{d\theta}{dt} = \frac{r_0^2\,\omega_0}{(r_0 + v_{rr}\,t)^2} \tag{30}$$

We can now integrate this to find $\theta(t)$. Again we need an initial condition. If

$$\theta(0) = \theta_0 \tag{31}$$

then

$$\theta(t) = \theta_0 + \int_0^t dt'\,\frac{r_0^2\,\omega_0}{(r_0 + v_{rr}\,t')^2} = \theta_0 + \frac{\omega_0\,r_0\,t}{r_0 + v_{rr}\,t} \tag{32}$$

There is another way of thinking about this which is slightly amusing. If I am a one dimensional creature living on this expanding ring, the variable I care about is not the angle, but the actual tangential distance I have traveled along the ring, which is

$$\ell(t) = \int_0^t r(t')\,\dot{\theta}(t')\,dt' \tag{33}$$

so that differentiating to get the velocity gives just the tangential velocity

$$\dot{\ell}(t) = r(t)\,\dot{\theta}(t) \tag{34}$$

Differentiating again to get the acceleration gives

$$\ddot{\ell}(t) = \frac{d}{dt}\Big(r(t)\,\dot{\theta}(t)\Big) = \frac{d}{dt}\Big(r(t)^2\,\dot{\theta}(t)/r(t)\Big) \tag{35}$$

But from (26), the product of the first two terms is a constant so

$$\ddot{\ell}(t) = r(t)^2\,\dot{\theta}(t)\,\frac{d}{dt}\Big(1/r(t)\Big) = r(t)^2\,\dot{\theta}(t)\,\frac{-\dot{r}(t)}{r(t)^2} = r(t)\,\dot{\theta}(t)\,\frac{-\dot{r}(t)}{r(t)} = -\frac{\dot{r}(t)}{r(t)}\,\dot{\ell}(t) \tag{36}$$

or

$$m\ddot{\ell}(t) = -m\frac{\dot{r}(t)}{r(t)}\,\dot{\ell}(t) \tag{37}$$

Now the point is that this looks like $F = ma$ for a particle subject to a velocity dependent frictional force,

$$-m\frac{\dot{r}(t)}{r(t)}\,\dot{\ell} \tag{38}$$

Creatures living on this expanding ring would feel this friction as a result of the expansion of their "space." It is very different from the frictional forces that we have talked about so far because this is a conservative system. No energy is being lost to heat, but instead, something that looks like friction is generated by the time dependence of the system. Something very much like this is going on in our expanding universe - it is called "Hubble friction." You see here that is it just conservation of a momentum (in this case angular) in the expanding space.

This system is animated in the *Mathematica* file **lecture-7-1.nb**.

Example - bead on a rotating rod

Next, consider a bead with mass m on a straight frictionless rod that rotates with constant angular velocity ω. For simplicity, we will let it rotate around the origin in the x-y plane. Then gravity plays no role. The angle with the x axis at time t is ωt. This would not be a trivial problem if we were armed only with $\vec{F} = m\vec{a}$. But with a Lagrangian it is easy. For one thing, the Lagrangian technique allows us to focus just on where the bead is on the wire. Let ℓ (which can be negative) be the distance along the rod where the bead sits at time t. This specifies the configuration of the system. The position of the bead at time t is then

$$x = \ell\cos\omega t \qquad y = \ell\sin\omega t \tag{39}$$

The velocity is

$$\dot{x} = \dot{\ell}\cos\omega t - \ell\omega\sin\omega t \qquad \dot{y} = \dot{\ell}\sin\omega t + \ell\omega\cos\omega t \tag{40}$$

Thus the kinetic energy is

$$\frac{1}{2}m\Big(\dot{x}^2 + \dot{y}^2\Big) = \frac{1}{2}m\Big(\dot{\ell}^2 + \ell^2\omega^2\Big) \tag{41}$$

The explicit time dependence from (40) has gone away when we form the square of velocity. If you think about the two terms on the right hand side of (41), you will realize that the first one is related to the motion of the bead along the rod, and the second one with the rotation of the rod. Because these two motions are perpendicular, there is no cross term and the total kinetic energy is just a sum of the two effects.

In this case, there is no potential, so the Lagrangian is just given by the kinetic energy, (41), and it doesn't depend on t explicitly, so we don't need t in the list of variables,

$$\mathcal{L}(\ell, \dot{\ell}) = \frac{1}{2} m \Big(\dot{\ell}^2 + \ell^2 \omega^2 \Big) \tag{42}$$

and the Euler-Lagrange equation is

$$0 = m \ell \omega^2 - \frac{d}{dt}(m \dot{\ell}) \tag{43}$$

or

$$\ddot{\ell} = \omega^2 \ell \tag{44}$$

This is a differential equation we can solve easily because it is linear and TTI. The general solution is

$$\ell(t) = C\, e^{\omega t} + D\, e^{-\omega t} \tag{45}$$

You are likely to see this general solution written as

$$\ell(t) = A \cosh \omega t + B \sinh \omega t \tag{46}$$

in term of the so-called hyberbolic functions,

$$\sinh x = \frac{e^x - e^{-x}}{2} \qquad \cosh x = \frac{e^x + e^{-x}}{2} \tag{47}$$

These are handy, because they have properties that are reminiscent of the more familiar trigonometic functions, $\sin x$ and $\cos x$. But if you are not used to them, you can always do everything in terms of exponentials. At any rate, if the bead starts from rest at $\ell = \ell_0$ at $t = 0$, the solution looks like

$$\ell(t) = \ell_0 \cosh \omega t = \ell_0 \, \frac{e^{\omega t} + e^{-\omega t}}{2} \tag{48}$$

Exponentials get big quickly, so this bead gets going pretty fast. For example, in the time π/ω, while the rod rotates through 180°, the distance from the origin increases by a factor of

$$\cosh \pi \approx 11.6 \tag{49}$$

So this is at good way to launch things.

If instead, the rod rotates with constant angular velocity ω around the origin in the x-z plane in the earth's gravitational field, then we have to include the effect of gravity - but this is easy. Again, we take the angle with the x axis at time t is ωt.

The configuration of the system looks the same except that z replaces y. Thus the kinetic energy looks the same. But now there is a gravitational contribution to the potential energy

$$U(\ell, t) = mgz = mg\ell \sin \omega t \tag{50}$$

Thus the Lagrangian in this case depends explicitly on t. It looks like

$$\mathcal{L}(\ell, \dot{\ell}, t) = \frac{1}{2} m \Big(\dot{\ell}^2 + \ell^2 \omega^2 \Big) - mg\ell \sin \omega t \tag{51}$$

and the Euler-Lagrange equation is

$$0 = m\ell\omega^2 - mg \sin \omega t - \frac{d}{dt}(m\dot{\ell}) \tag{52}$$

or

$$\ddot{\ell} = \omega^2 \ell - g \sin \omega t \tag{53}$$

More degrees of freedom

The Lagrangian works just as well for more particles, or in more dimensions. Suppose that there are n particles, so that

$$T = \sum_{j=1}^{n} \frac{m_j}{2} \dot{x}_j^2 \quad \text{and} \quad U = V(x_1, \cdots, x_n) \tag{54}$$

Then it is easy to see by the same sort of arguments that

$$\frac{\delta S}{\delta x_j} = -m_j \, \ddot{x}_j - V_j \Big(x_1, \cdots, x_n \Big) \tag{55}$$

where

$$-V_j(x_1, \cdots, x_n) \equiv -\frac{\partial}{\partial x_j} V(x_1, \cdots, x_n) \tag{56}$$

is the force on particle j. For $S[x]$ to be an extremum, we must have (55) vanish for each j, which just gives $F = ma$ for each particle.

If the Lagrangian depends on more degrees of freedom, there is an Euler-Lagrange equation that must be satisfied for each coordinate:

$$\frac{\partial}{\partial x_j} \mathcal{L}\Big(x, \dot{x}, t\Big) - \frac{d}{dt} \frac{\partial}{\partial \dot{x}_j} \mathcal{L}\Big(x, \dot{x}, t\Big) = 0 \tag{57}$$

where the xs and $\dot{x}$s in the function $\mathcal{L}$ now indicate all the components — that is $\mathcal{L}(x, \dot{x}, t)$ is really a shorthand for

$$\mathcal{L}(x_1, x_2, \cdots, \dot{x}_1, \dot{x}_2, \cdots, t) \tag{58}$$

lecture 8

Topics:
- Why Euler-Lagrange equations?
- Hamilton's principle
- Functions of functions
- Calculus of variations
- Functional derivatives
- Finding functional derivatives
- Back to Hamilton's principle
- The Lagrangian and the action
- Quantum mechanics and the classical trajectory
- Appendix: On the functional Taylor series

Why Euler-Lagrange equations?

In the last week, you have done the really important part of your study of Lagrangians, at least for this course. You have have learned how to construct them and to derive from them the Euler-Lagrange equations that replace $F = ma$ as our equations of motion. You have seen in a non-trivial example (the second problem on as3) that the physical results are independent of the particular choice of coordinates you use to write down the Lagrangian. Next time, we will see how this gives a beautiful explanation of the great conservation laws of energy and momentum in terms of fundamental symmetries of the world. I hope that all of this has convinced you that the Lagrangian formulation of mechanics really simplifies your life. We won't come close in this course to seeing all the advantages of this beautiful way of doing mechanics (you can get a little sense of it by reading though the sections of Morin that I have not assigned). You can see more in Physics 151 if your appetite is whetted.

Nevertheless, as I expected, and indeed as I saw on the QA this week, most of you would kind of like to understand why it works! This is not **really** fair of course. You don't know why $F = ma$ works either. You are just so used to it that it seems reasonable. But there is something that we can say about that, while it is hard, will help you understand some of the magic of Lagrangian. And, as I will try to convince you, it is telling something about the way the world really works at a deeper level.

To do this, I have to introduce some mathematics that will probably be hard for most you - essentially the calculus of an infinite number of variables! This probably sounds really scary, and it is scary. **But DON'T PANIC!** Remeber that you have already done the really important part. This next week is mostly for your general education. But you will find, I hope, that this deeper formulation not only increases your understanding, but will give you a useful bag of tricks for dealing with practical problems.

Here is the deal. The Euler-Lagrange equations can be derived from something called "Hamilton's principle" which is the statement (which we will make more carefully in a moment) that the classical trajectory is the path of the classical system through time for which small changes in the path do not change a quantity called the "Action" which is obtained by integrating the Lagragian over the path. Most of todays class will be taken you by trying to understand what this statement means and showing that it give rise to the Euler-Lagrange equations. At the end, I will try to give you at least a hand-waving idea of how this is related to quantum mechanics.

Hamilton's principle

We will start with the very simplest example - a mass moving in a potential in one dimension. Consider the one dimensional motion of a particle with coordinate x in a potential $V(x)$. Call the particle's kinetic energy

$$T(\dot{x}) = \frac{m}{2}\dot{x}^2 \tag{1}$$

Now we would like to find a particular trajectory $x(t)$ are such that the particle moves from x_1 at time t_1 to x_2 at time t_2 — or

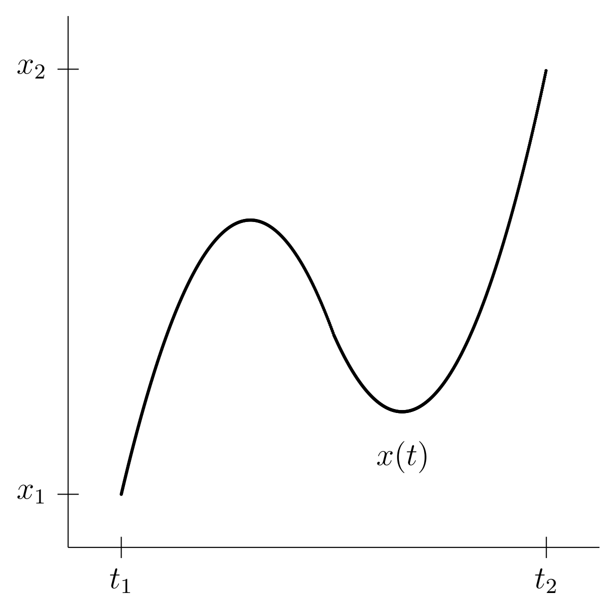
$$x(t_1) = x_1 \qquad x(t_2) = x_2 \tag{2}$$

Hamilton's principle is the statement that if you compute the quantity $S[x]$ (called the **action** — I'm not really sure why) - the following integral of the Lagrangian:

$$S[x] \equiv \int_{t_1}^{t_2} \Big[\mathcal{L}\big(\dot{x}(t)\big)\Big]\, dt = \int_{t_1}^{t_2} \Big[T\big(\dot{x}(t)\big) - V\big(x(t)\big)\Big]\, dt \tag{3}$$

depending on a function $x(t)$ satisfying[1] then the variation of $S[x]$ with respect to the function $x(t)$ vanishes for the actual trajectory. The hard part of Hamilton's principle will not be proving this statement, but figuring out exactly what it means to talk about the variation of something with respect to a function.

The first thing to notice is that the Lagrangian $\mathcal{L} = T - V$ is an ordinary function of two variables, x and $\dot{x}$,[2] but that $S[x]$ is actually a function of a function — it depends on $x(t)$ for all values of t from t_1 to t_2 where $x(t)$ can be any function satisfying (2) — something like



[1] We will see later that the restriction to functions satisfying (2) is important and has to do with the initial conditions that we have spent so much time talking about in the first couple of weeks.

[2] In the simple example we are working out now, T is a function of one of them and V of the other, but in general, both T and V could depend on both x and $\dot{x}$.

Thus $S[x]$ is a function whose argument is itself a function (such an object is sometimes called a **functional** and I may sometimes use that term). I have put the x in square brackets to remind you that in this case x is a function rather than a number. The value of $S[x]$, on the other hand, is just a number. It doesn't depend on t. The variable t in (3) is a dummy variable. Now Hamilton's principle is a statement about the variation of $S[x]$ as we let $x(t)$ vary over all possible functions satisfying (2).[3] We will find that when $x(t)$ is very near to a solution to Newton's second law, then $S[x]$ changes slowly as a function of whatever parameters you use to specify the function $x(t)$. When $x(t)$ is a solution, $S[x]$ is not varying at all.

Functions of functions

I imagine that the idea of a functional — that is a function of a function — is pretty unfamiliar to most of you, but you shouldn't get too worried about it. Most of you have learned or are beginning to learn about the calculus of functions of several variables, and dealing with functionals is not really much different. The main differences are in notation. In fact, I would argue that the big jump is going from one or two variables to more than that. Once you have made it to three variables, it really doesn't get any harder to think about more — not even an infinite number more.

Here is what I mean by the peculiar statement that going from two variables to three variables is what is hard. This has to do with visualization. As you have probably noticed, it gets harder to make a mental picture of what a function means as the number of variables increase. A function of a single variable is easy. We naturally associate a function $f(x)$ with a graph of $y = f(x)$. This allows us to do things that bring the function alive to us, such as relating the derivative of the function to the slope of the line in the graph. We can sort of do something similar with a function of two variables $f(x, y)$ by imagining a surface in three dimensional space with $z = f(x, y)$. This can be quite helpful, because it allows you to have a visual representation of features that only appear with more than one variable, such as the gradiant vector, which for two variables x and y looks like the two dimensional vector

$$\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) \tag{4}$$

The vector $\vec{\nabla} f(x, y)$ points in a direction that corresponds to going upwards on the surface $z = f(x, y)$.

But what do you do about visualizing functions of three variables, or more? There is really no ideal way of doing this. So already by the time you get to functions of three variables, you have to stop relying on visual crutches and just develop an analytic sense of what the function means.

A functional like $S[x]$ is just a function of an infinite number of variables where the variables are the values of $x(t)$ at all possible values of t. The really peculiar new thing about a functional is that the variables are labeled by a continuous parameter, rather than having different names like x, y and z, or different discrete indices, like a_1, a_2, *etc*. This is why functionals look so different and why we have to invent some new notation to deal with them.

[3]Really? All possible functions? One has to be reasonable here — it might be better to say all possible functions for which the integral in (3) makes sense — differentiable functions in this example. We will ignore such issues and leave them for the mathematicians.

A mathematical example of a functional is the length of a path described by a curve, $y = f(x)$. The path length from x_1 to x_2 is

$$P[f] = \int_{x_1}^{x_2} \sqrt{dx^2 + dy^2} = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx = \int_{x_1}^{x_2} \sqrt{1 + f'(x)^2}\, dx \tag{5}$$

The path length depends on the function f that defines the shape of the curve. This example is particularly interesting because it depends on the endpoints in same way that (3) does.

A specific example of path length may be useful. Consider a graph $y = f(x)$ for the function

$$f(x) = \sqrt{R^2 - x^2} \tag{6}$$

$$f'(x) = -x/\sqrt{R^2 - x^2} \tag{7}$$

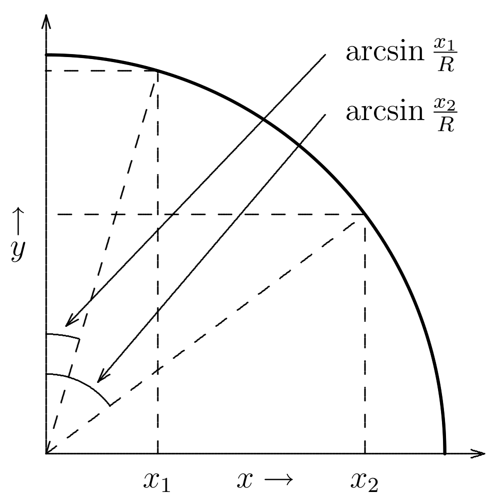
Then the path length from x_1 to x_2 is

$$\int_{x_1}^{x_2} \sqrt{1 + f'(x)^2}\, dx = \int_{x_1}^{x_2} R/\sqrt{R^2 - x^2}\, dx \tag{8}$$

which with the substitution $x = R \sin\theta$ becomes

$$R \int_{\arcsin x_1/R}^{\arcsin x_2/R} d\theta = R \left(\arcsin \frac{x_2}{R} - \arcsin \frac{x_1}{R}\right) \tag{9}$$

And this is right because this is just the arc length along a circle, as show in the diagram below.



Here is a more physical example. Suppose that the function $y = f(x)$ in (5) describes the height of the track of of roller coaster. If a roller car is moving along the track with mass m and energy

$$E = \frac{1}{2}mv^2 + mgy = \frac{1}{2}mv^2 + mg\, f(x) \tag{10}$$

the time the car takes to get from x to $x + dx$ is the infinitesimal path length, $\sqrt{dx^2 + dy^2} = dx\, \sqrt{1 + f'(x)^2}$ divided by the speed v which from (10) is

$$v = \sqrt{2\big(E - mg\, f(x)\big)/m} \tag{11}$$

Thus the time the car takes to get from x_1 to x_2 is

$$\tau[f] = \int_{x_1}^{x_2} \frac{\sqrt{dx^2+dy^2}}{v} = \int_{x_1}^{x_2} \frac{\sqrt{1+f'(x)^2}}{\sqrt{2\Big(E - mg\, f(x)\Big)/m}}\, dx \tag{12}$$

The time (12) is a function of E, m and g, but it is a **functional** of the function $f(x)$ that describes the height of the track. This is an important example to understand and we will try to come back to it several times over the next few weeks.

Calculus of variations

Now let's return to Hamilton's principle. The proof and indeed the more precise mathematical statement of the principle is an exercise in what is called the calculus of variations, which is sort of calculus with an infinite number of variables. It probably won't surprise you that what we need to do is to generalize the notion of a derivative to functionals and to explore the nature of the Taylor series. After all, these are our main tools in understanding how things vary.

The actual example we are interested in is complicated because the integrand on the right hand side of (3) depends on both $x(t)$ and $\dot{x}(t)$, so let us first discuss a couple of simpler examples.

Consider the quantity

$$W[x] = \int_{t_1}^{t_2} \Big(x(t) - vt\Big)^2\, dt \tag{13}$$

for some constant v and find the value of $x(t)$ for which the variation of $W[x]$ with respect to $x(t)$ vanishes. We can do this by the following trick. Suppose that we have already found the soltuion $x(t)$. Now consider small variations about this particular function:

$$x(t) + \delta x(t) \tag{14}$$

I've used the symbol δ rather than Δ just to remind us that this is a "small" function rather than a small number. Now look at (13) as function of $x + \delta x$,

$$\begin{gathered} W[x+\delta x] = \int_{t_1}^{t_2} \Big(x(t) + \delta x(t) - vt\Big)^2\, dt \\ = \int_{t_1}^{t_2} \Big(x(t) - vt\Big)^2\, dt + \int_{t_1}^{t_2} \delta x(t)\, 2\Big(x(t) - vt\Big)\, dt + \int_{t_1}^{t_2} \Big(\delta x(t)\Big)^2\, dt \\ = W[x] + \int_{t_1}^{t_2} \delta x(t)\, 2\Big(x(t) - vt\Big)\, dt + \mathcal{O}(\delta x^2) \end{gathered} \tag{15}$$

Now we look at the coefficient of the linear term in $\delta x(t)$, which is $2\Big(x(t) - vt\Big)$. This vanishes **for all** possible small functions $\delta x(t)$ if

$$x(t) = vt \tag{16}$$

and this is the answer we are looking for. Around the function $x(t) = vt$, there is no linear term in the expansion of $W[x]$. Note that while t in (15) is a dummy variable, we can say something about the t dependence because we require that the variation vanishes for any possible $\delta x(t)$. We will come back to this below when discuss functional derivatives.

This result is quite reasonable and maybe obvious in this particular example because $W[x]$ is actually minimized for $x(t) = vt$. The variation of a functional at a minimum vanishes for the same reason that the derivative of a smooth function vanishes at a minimum. If there were a linear variation, then on one side or the other, the result would be larger, which is impossible at a minimum value. But what we want to extract from this calculation is not the result, but the general technique.

Functional derivatives

What we did to solve the variational problem for (13) was to set to zero the coefficient of the linear term in $\delta x(t)$ in $W[x + \delta x]$. It is useful to give this coefficient a name, so we call it a **functional derivative**, and denote it by

$$\frac{\delta W}{\delta x(t)}[x] \quad \text{or} \quad \frac{\delta}{\delta x(t)} W[x] \tag{17}$$

so in this case, we can write

$$\frac{\delta W}{\delta x(t)}[x] = 2\Big(x(t) - vt\Big) \tag{18}$$

Notice that the t dependence of the RHS of (18) comes from the t dependence of the differential in the denominator of the LHS — again $W[x]$ for some particular x is just a number with no t dependence.

This definition of functional derivative is generally useful, so let's discuss it in general:

$$\frac{\delta}{\delta x(t)} W[x] \text{ is the coefficient of the linear term in } \delta x(t) \text{ in } W[x + \delta x]. \tag{19}$$

This is a reasonable definition because it makes the functional Taylor series work in the same way that the ordinary Taylor series does. In terms of the functional derivative, the functional Taylor series starts like this:

$$W[x + \delta x] = W[x] + \int_{t_1}^{t_2} \delta x(t) \, \frac{\delta}{\delta x(t)} \, W[x] \, dt + \cdots \tag{20}$$

In words, this says that the total change in the functional is obtained by adding up (actually integrating because the variable is continuous) the small changes $\delta x(t)$ in $x(t)$ at each point times the rate of change of the functional at that point. This is just what we always say, except that the number of variables has become continuously infinite. (20) is precisely analogous to the Taylor series for functions of several variables that we will discuss in the appendix and in future lectures:

$$F(x_1 + \Delta x_1, x_2 + \Delta x_2, \cdots) = F(x_1, x_2, \cdots) + \sum_j \Delta x_j \, \frac{\partial}{\partial x_j} F(x_1, x_2, \cdots) + \cdots \tag{21}$$

The only difference is that instead of the sum over all the different variables that we have in (21), in (20) we need an integral over t because in effect, the value of $x(t)$ at each value of t is a separate variable.

Finding functional derivatives

In general, functional derivatives can be difficult to find, but it is easy to find them for functionals like $W[x]$ that have the form of an integral of an ordinary function —

$$W[x] = \int_{t_1}^{t_2} F\Big(x(t)\Big)\,dt \tag{22}$$

In this case, the functional derivative of $W[x]$ is related to the ordinary derivative of F —

$$\frac{\delta W}{\delta x(t)}[x] = F'\Big(x(t)\Big) \tag{23}$$

The reason this is so simple is that we can calculate the functional Taylor expansion of W using the ordinary Taylor expansion of F, and looking for the coefficient of $\delta x(t)$. Let's show how this works.

$$W[x+\delta x] = \int_{t_1}^{t_2} F\Big(x(t)+\delta x(t)\Big)\,dt \tag{24}$$

But using the ordinary Taylor expansion for the ordinary function $F(x)$, the right hand side becomes

$$= \int_{t_1}^{t_2} \Big[F\Big(x(t)\Big) + \delta x(t)\,F'\Big(x(t)\Big) + \cdots\Big]\,dt \tag{25}$$

Picking out the coefficient of $\delta x(t)$ on the right hand side gives (23). Notice the way that the variable t gets promoted from being a dummy variable in (22) and (24) to being a real variable in (25). We single out a particular value of t when we find the coefficient of $\delta x(t)$ for that t.

The function F in (22) may depend on other functions of t, in which case we can generalize (22) as follows:

$$\frac{\delta}{\delta x(t)} \int_{t_1}^{t_2} F\Big(x(t'), y(t'), \cdots\Big)\,dt' = \frac{\partial}{\partial x}\,F(x, y(t), \cdots)\big|_{x=x(t)} \tag{26}$$

where ... denotes and other functions that F depends on. Note that in (26), t' is a dummy index, integrated over so that it could be called anything except t. The dependence on t is real (not dummy) because the process of functional differentiation picks out a value of t. Here are some examples:

$$\begin{aligned}
F[x] = \int_{t_1}^{t_2} x(t')^3\,dt' \quad &\Rightarrow \quad \frac{\delta}{\delta x(t)}F[x] = 3\,x(t)^2 \\
F[x] = \int_{t_1}^{t_2} \sin\Big(x(t')\Big)\,dt' \quad &\Rightarrow \quad \frac{\delta}{\delta x(t)}F[x] = \cos\Big(x(t)\Big) \\
F[x,y] = \int_{t_1}^{t_2} \Big(x(t')\,y(t')\Big)^3\,dt' \quad &\Rightarrow \quad \frac{\delta}{\delta x(t)}F[x,y] = 3\,x(t)^2 y(t)^3 \\
F[x,y] = \int_{t_1}^{t_2} \sin\Big(x(t')\,y(t')\Big)\,dt' \quad &\Rightarrow \quad \frac{\delta}{\delta x(t)}F[x,y] = y(t)\cos\Big(x(t)\,y(t)\Big)
\end{aligned} \tag{27}$$

We can summarize these ideas by saying that when the functional is the integral of an ordinary function, what the functional derivative does is to eliminate the integral and differentiate the function. Think about it in the context of the examples in (27).

Back to Hamilton's principle

With this new tool of the calculus of variations, we can go back and consider Hamilton's principle. Remember that what we want to do is to show that the functional $S[x]$ has zero variation around the same trajectory that we get by using $F = ma$.

According to our discussion, what we want to do to impose vanishing variation on $S[x]$ is to set the functional derivative of $S[x]$,

$$\frac{\delta S}{\delta x(t)} \tag{28}$$

to zero. As in the simpler example, we calculate the functional derivative by performing a functional Taylor series and picking out the coefficient of $\delta x(t)$. For pedagogical purposes, we will break this up into two pieces:

$$S = S_T - S_V \tag{29}$$

where

$$S_T[x] \equiv \int_{t_1}^{t_2} \left[T\Big(\dot{x}(t)\Big)\right] dt = \int_{t_1}^{t_2} \frac{m}{2}\Big(\dot{x}(t)\Big)^2 dt \tag{30}$$

and

$$S_V[x] \equiv \int_{t_1}^{t_2} \left[V\Big(x(t)\Big)\right] dt = \int_{t_1}^{t_2} V\Big(x(t)\Big)\, dt \tag{31}$$

For the S_V term, because V does not depend on $\dot{x}$, this is just like the $W[x]$ example,

$$\frac{\delta S_V}{\delta x(t)} = V'\Big(x(t)\Big) \tag{32}$$

This is a good sign because the derivative of the potential is related to the force.

For the S_V, term, we will need an extra step, so let's write out the functional Taylor series in detail.

$$S_T[x + \delta x] = \int_{t_1}^{t_2} \frac{m}{2}\Big(\dot{x}(t) + \delta\dot{x}(t)\Big)^2 dt \tag{33}$$

$$= \int_{t_1}^{t_2} \frac{m}{2}\Big(\dot{x}(t)\Big)^2 dt + \int_{t_1}^{t_2} \delta\dot{x}(t)\, m\, \dot{x}(t)\, dt + \mathcal{O}(\delta x^2) \tag{34}$$

$$= S_T[x] + \int_{t_1}^{t_2} \delta\dot{x}(t)\, m\, \dot{x}(t)\, dt + \mathcal{O}(\delta x^2) \tag{35}$$

But the $\delta\dot{x}(t)$ in (35) is the change in $\dot{x}(t)$ when we make a change $\delta x(t)$ in $x(t)$. Thus since

$$\frac{d}{dt}\Big(x(t) + \delta x(t)\Big) = \dot{x}(t) + \frac{d}{dt}\delta x(t) \tag{36}$$

$\delta\dot{x}(t)$ is given by

$$\delta\dot{x}(t) = \frac{d}{dt}\delta x(t)\,. \tag{37}$$

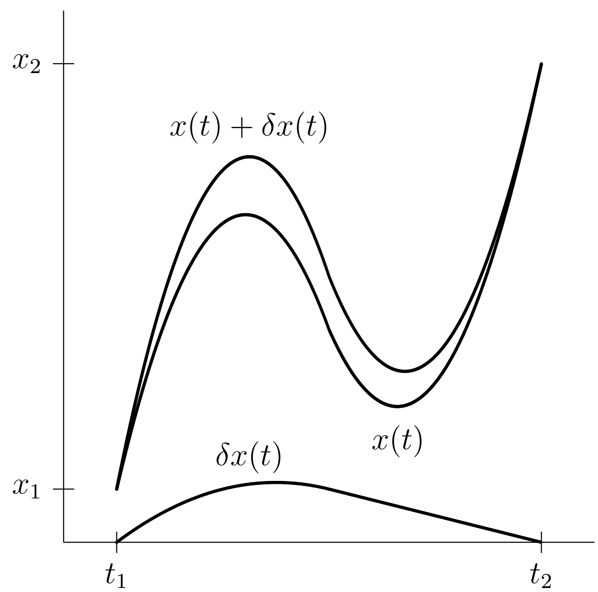
The trouble with (35) is that it depends on $\delta\dot{x}(t)$, while we have defined the functional derivative to be the coefficient of $\delta x(t)$. To take care of that, we can integrate by parts:

$$\begin{aligned} \int_{t_1}^{t_2} \delta\dot{x}(t)\, m\, \dot{x}(t)\, dt &= \int_{t_1}^{t_2} \left(\frac{d}{dt}\delta x(t)\right) m\, \dot{x}(t)\, dt \\ &= \int_{t_1}^{t_2} \frac{d}{dt}\Big(\delta x(t)\, m\, \dot{x}(t)\Big)\, dt - \int_{t_1}^{t_2} \delta x(t)\, m\, \ddot{x}(t)\, dt \\ &= -\int_{t_1}^{t_2} \delta x(t)\, m\, \ddot{x}(t)\, dt + \delta x(t_2)\, m\, \dot{x}(t_2) - \delta x(t_1)\, m\, \dot{x}(t_1) \end{aligned} \tag{38}$$

The last line in (38) has two terms from the endpoints of the integration. These don't look like everything else, and we want to get rid of them. This is where (2) comes in. We are supposed to be restricting our attention to functions satisfying (2). But if $x(t)$ and $x(t) + \delta x(t)$ both satisfy (2), then

$$\delta x(t_1) = \delta x(t_2) = 0\,. \tag{39}$$

In words, we are only interested in variations that vanish at the endpoints. These functions might look something like this



So in fact, the funny looking terms vanish because of (2). Then (38) becomes

$$\int_{t_1}^{t_2} \delta\dot{x}(t)\, m\, \dot{x}(t)\, dt = -\int_{t_1}^{t_2} \delta x(t)\, m\, \ddot{x}(t)\, dt \tag{40}$$

Now we can pick out the coefficient of $\delta x(t)$ and see that

$$\frac{\delta S_T}{\delta x(t)} = -m\, \ddot{x}(t) \tag{41}$$

Now, finally, we can verify that Hamilton's principle is correct. Putting (32) and (41) together, we have that the condition for vanishing variation of $S[x]$ is

$$\frac{\delta S}{\delta x(t)} = -m\, \ddot{x}(t) - V'\Big(x(t)\Big) = 0 \tag{42}$$

or

$$m\,\ddot{x}(t) = -V'\Big(x(t)\Big) \tag{43}$$

which, since $F = -V'$, is just Newton's second law, which in this case, we know, is also the Euler-Lagrange equation.

The Lagrangian and the action

As we will see, Hamilton's principle really captures more of what is going on in the world than $F = ma$. Let's see consequences of Hamilton's principle again in a more general language in terms of the **Lagrangian**, $\mathcal{L}$. Again the **action** is

$$S[x] = \int_{t_1}^{t_2} \mathcal{L}\Big(x(t), \dot{x}(t)\Big)\, dt \tag{44}$$

Here, x might actually have indices that allows it to represent more than one particle or dimension or both. We won't write them explicitly in this formal derivation. Now we want the functional derivative of $S[x]$ with respect to $x(t)$ to vanish. In general, the functional derivative is

$$\frac{\delta S}{\delta x(t)}[x] = \frac{\partial}{\partial x(t)}\mathcal{L}\Big(x(t), \dot{x}(t)\Big) - \frac{d}{dt}\frac{\partial}{\partial \dot{x}(t)}\mathcal{L}\Big(x(t), \dot{x}(t)\Big) \tag{45}$$

The first term arises from the Taylor expansion of the $x(t)$ dependence. The second term arises in the same way as (41), from the Taylor expansion of the $\dot{x}(t)$ dependence, followed by an integration by parts, which gives the minus sign. Thus Hamilton's principle implies that the solution for the motion satisfies the Euler-Lagrange equation(s)

$$\frac{\partial}{\partial x(t)}\mathcal{L}\Big(x(t), \dot{x}(t)\Big) - \frac{d}{dt}\frac{\partial}{\partial \dot{x}(t)}\mathcal{L}\Big(x(t), \dot{x}(t)\Big) = 0 \tag{46}$$

If x has several components, (46) must be true for each component separately.

Quantum mechanics and the classical trajectory

In my view, the most important reason that Hamilton's principle and the Lagrangian are so important has to do not with classical mechanics alone, but with quantum mechanics and the way in which classical physics emerges as an approximation to the quantum world.

From the classical point of view, Hamilton's principle is actually a little peculiar. Why should it matter to a classical particle what the value of the action is for paths that the particle does not actually take? OK – - so Hamilton's principle works to give the classical equation of motion, but it is hard to figure out what it means physically. But in quantum mechanics, it has a very definite meaning, because in quantum mechanics, the particle really takes all paths! This is absolutely nutty, but this is really the way the world works. Roughly, the way it works is this. When a quantum mechanical particle moves from point x_1 at time t_1 to point x_2 at time t_2, it takes all trajectories from the starting point to the end simultaneously. But the different trajectories can add together like the different ripples in a wave on a pond. Associated with each trajectory there is a complex number A whose phase is the action divided by $\hbar$, Planck's constant over 2π.

$$e^{iS[x]/\hbar} \tag{47}$$

The most likely trajectories are those that are near the classical trajectory, because the action is changing very slowly for these trajectories, they have approximately the action of the classical trajectory, and therefore all the As have the same phase and all the nearby trajectories add up coherently. Trajectories very far from the classical trajectory are unlikely because the phase is changing rapidly and nearby trajectories have different phases and add nearly to zero.

There is another way of putting this that is perhaps more interesting. Euler-Lagrange equations are differential equation, but not every differential equation is an Euler-Lagrange equation. The fact that the world is quantum mechanical explains why the classical physics that we see can be described by the solutions not just to any old differential equation, but specifically to Euler-Lagrange equations. We will see why this distinction is important as we go along.

Appendix: On the functional Taylor series

It might be useful for some of you if I expand on the statement of the functional Taylor series because it a pretty piece of math. Some of you, on the other hand, may find this rather terrifying. Please DON'T PANIC. This section is called an **appendix** because it is completely optional. I won't get to it at all in lecture. But for some of you it may be fun.

The picture of the functional Taylor series that I will give you is related to the following statement that you may find useful. The t in a function $x(t)$ can be thought of as a kind of index, labeling components of an infinite dimensional vector. Or to put this the other way around, a vector $\vec{r}$ can be thought of as a function of the index that labels the component — r_j is a number for each j just like $x(t)$ is a number for each t. In this way of thinking, a functional, $W[x]$ is like a function of several variables, $f(r_1, \cdots, r_n)$ (depending on an n-dimensional "vector"), except that $W[x]$ depends on an infinite number of variables, the values of $x(t)$ for some range of t. Now what does the Taylor expansion look like for a function of several variables, like $f(r_1, \cdots, r_n)$? We can build up the Taylor series by looking at one variable at a time:

$$
\begin{aligned}
f(r_1 + a_1, \cdots, r_n) &= \left(1 + a_1 \frac{\partial}{\partial r_1} + \frac{1}{2} a_1^2 \frac{\partial^2}{\partial r_1^2} + \cdots\right) f(r_1, \cdots, r_n) \\
f(r_1 + a_1, r_2 + a_2, \cdots, r_n) &= \left(1 + a_2 \frac{\partial}{\partial r_2} + \frac{1}{2} a_2^2 \frac{\partial^2}{\partial r_2^2} + \cdots\right) f(r_1 + a_1, r_2, \cdots, r_n) \\
& \text{and so on for all } n \text{ variables}
\end{aligned}
\tag{48}
$$

This looks complicated, but it can be simplified easily using a very beautiful form for the Taylor expansion of one variable:

$$
g(x + a) = \exp\left(a \frac{\partial}{\partial x}\right) g(x) = \sum_{k=0}^{\infty} \frac{a^k}{k!} \frac{d^k}{dx^k} g(x) \tag{49}
$$

The expansion of the exponential precisely reproduces the terms in the Taylor expansion. Using (49), equation (48) becomes

$$
f(r_1 + a_1, \cdots, r_n + a_n) = \left(\prod_{j=1}^{n} \exp\left(a_j \frac{\partial}{\partial x_j}\right)\right) f(r_1, \cdots, r_n) \tag{50}
$$

But because the product of exponentials is the exponential of the sum of the exponents, this can be written as

$$f(r_1 + a_1, \cdots, r_n + a_n) = \exp\left(\sum_{j=1}^{n} a_j \frac{\partial}{\partial x_j}\right) f(r_1, \cdots, r_n) \tag{51}$$

This is the Taylor series for more than one variable. It is rather neat that it looks just like the Taylor series for a single variable, (49), except that we have to include derivatives with respect to all the variables in the exponent. Now for an infinite number of variables labeled by a continuous variable t, in some range $t_1 \leq t \leq t_2$, the argument goes the same way, but instead of summing over the indices, as we do in (51), we must integrate over the continuous variable t, so that the functional Taylor series looks like

$$W[x + \delta x] = \exp\left(\int_{t_1}^{t_2} dt\, \delta x(t)\, \frac{\delta}{\delta x(t)}\, dt\right) W[x] \tag{52}$$

Note that if we expand the exponential, we get a first term which is just (20), as expected from our earlier discussion.

lecture 9

Topics:

- Where are we now?
- Energy again
- Example - frictionless table
- When is F the energy?
- Symmetries and transformations
- Example: space translations for one particle
- Space translations for two particles
- Space translations for many particles
- Finding Symmetries
- Rotations
- Noether's theorem
- Momentum conservation from Noether's theorem
- More on rotations
- What functions are invariant?
- Example of functionals - Soap bubbles

Where are we now?

There is much more in these notes than we can possibly cover in class. But I thought I would put them all up and then decide what to discuss in detail after I see your QA responses tomorrow morning.

We have seen how conserved quantities can arise as the generalized momenta associated with variables that do not appear in the Lagrangian. Here we will discuss a generalization of this fact, that is an even more important principle. In Lagrangian mechanics, continuous symmetries lead to conserved quantities. We have already seen one example of this in our discussion of energy, and we will begin by making the connection with time translation invariance. We will go on to discuss symmetries more generally, and also in more detail the specific example of space translation symmetry, which leads to the conservation of total momentum. Because we will be talking about symmetry, at the risk of encouraging people to waste time, I have included a *Mathematica* file Kaleidoscope.nb that produces kaleidoscopic images with various different symmetries. I like it both because I like symmetry, and because it was so simple to construct with complex numbers (if only I could have worked in the Taylor expansion . . .).

Soap bubbles are a beautiful example of a functional minimization problem. The last sections discusses a simple type of soap bubble for which the analysis is the same as in Hamilton's principle.

Energy again

Now we are going to do a little math that will lead to a remarkable and beautiful result. If the Lagrangian does not explicitly depend on time, we will find that we can construct a function of the coordinates and the velocities of the system that does not change with time for any solution to the Euler-Lagrange equation. Often, we can identify this quantity with the energy. We will do this first for a single degree of freedom, and then extend the result to systems with more degrees of freedom.

Consider a system with a single degree of freedom described by the Lagrangian

$$\mathcal{L}(q, \dot{q}, t) \tag{1}$$

for some single coordinate q, and construct the quantity

$$F = \dot{q}\, \frac{\partial \mathcal{L}}{\partial \dot{q}} - \mathcal{L} \tag{2}$$

In general, F may depend on q, $\dot{q}$ and t. Let us now ask how F changes with time, by taking the total derivative

$$\frac{d}{dt}\, F = \frac{d}{dt} \left(\dot{q}\, \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) - \frac{d}{dt}\, \mathcal{L} \tag{3}$$

In the first term on the right hand side of (3), we use the product rule to write

$$\frac{d}{dt} \left(\dot{q}\, \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = \ddot{q}\, \frac{\partial \mathcal{L}}{\partial \dot{q}} + \dot{q}\, \frac{d}{dt}\, \frac{\partial \mathcal{L}}{\partial \dot{q}} \tag{4}$$

In the second term on the right hand side of (3), we use the fact that the t dependence of $\mathcal{L}$ comes from the explicit t dependence, and also from the implicit dependence on t through q and $\dot{q}$:

$$\frac{d}{dt}\, \mathcal{L} = \ddot{q}\, \frac{\partial \mathcal{L}}{\partial \dot{q}} + \dot{q}\, \frac{\partial \mathcal{L}}{\partial q} + \frac{\partial \mathcal{L}}{\partial t} \tag{5}$$

The relation (5) is an example of one of those multivariable calculus things we have talked about before that will make your eyes glaze over if you just stare at the symbols. But if you translate it into words, it makes perfect sense. It says that the total rate of change of F with t is the rate of change of $\dot{q}$ times the rate at which F changes with $\dot{q}$ plus the rate of change of q times the rate at which F changes with q plus the rate of change from the explicit time dependence. It is simply a matter of adding up all the possible sources of time variation of the function F. Subtracting (5) from (4) and using the Lagrange equation

$$\frac{d}{dt}\, \frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial \mathcal{L}}{\partial q} \tag{6}$$

we get

$$\frac{d}{dt}\, F = -\frac{\partial \mathcal{L}}{\partial t} \tag{7}$$

Thus if $\mathcal{L}$ does not depend on time EXPLICITLY, but only implicitly through the time dependence of q and $\dot{q}$, the function F is constant for the trajectory.

It is important to understand what is meant here by the words “explicit” and “implicit”. Explicit time dependence occurs only if there is some physics in the problem that changes with time. On the other hand, ANY function of q and $\dot{q}$ depends implicitly on time, because q and $\dot{q}$ for the trajectory depend on time.

Usually, this function F is the Energy! For example, suppose we look at the Lagrangian for a particle moving in a potential

$$\mathcal{L}(x, \dot{x}) = \frac{m}{2}\,\dot{x}^2 - V(x) \tag{8}$$

For this Lagrangian, the function F is

$$\dot{x}\,\frac{\partial}{\partial \dot{x}}\mathcal{L}(x,\dot{x}) - \mathcal{L}(x,\dot{x}) = \dot{x}\,m\,\dot{x} - \mathcal{L}(x,\dot{x}) = \frac{m}{2}\,\dot{x}^2 + V(x) \tag{9}$$

which is the energy, as promised.

For example, in the example we discussed last time of the bead on the horizontally rotating rod, where the Lagrangian is

$$\mathcal{L}(\ell, \dot{\ell}) = \frac{1}{2}m\Big(\dot{\ell}^2 + \ell^2\omega^2\Big) \tag{10}$$

the construction of (2) gives

$$F = \dot{\ell}\,m\dot{\ell} - \frac{1}{2}m\Big(\dot{\ell}^2 + \ell^2\omega^2\Big) = \frac{1}{2}m\Big(\dot{\ell}^2 - \ell^2\omega^2\Big) \tag{11}$$

These two examples are rather different. In the second, the fact that the Lagrangian has no explicit time dependence is an accident, arising from the cancellation of the ωt dependence that we talked about earlier. This is related to the fact that F in this case is not the kinetic plus the potential energy, but rather a curious combinations of the terms in the kinetic energy. It is conserved, but the physical interpretation is obscure.

In the first example, however, the particle moving in a potential, the Lagrangian does not depend explicitly on t because there is a symmetry of the system. The symmetry in this case is time translation invariance. In this case, F really is the physical energy. This is the first of several important examples we will see in this course of the connection between a symmetry and conservation law. We will explore this further next week.

For a vertically rotating rod, $\mathcal{L}$ depends on time explicitly because of the factor of $\sin\omega t$ in the potential energy, F is not conserved.

The construction of the energy function makes it clear how to deal with more degrees of freedom. If q has an index, q_j where j goes from 1 to n, the analogous construction for the function F is

$$F = \sum_j \dot{q}_j\,\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \mathcal{L} \tag{12}$$

It must have this form for the analog of (3)-(7) to be valid, because the analog of (5) for more degrees of freedom is

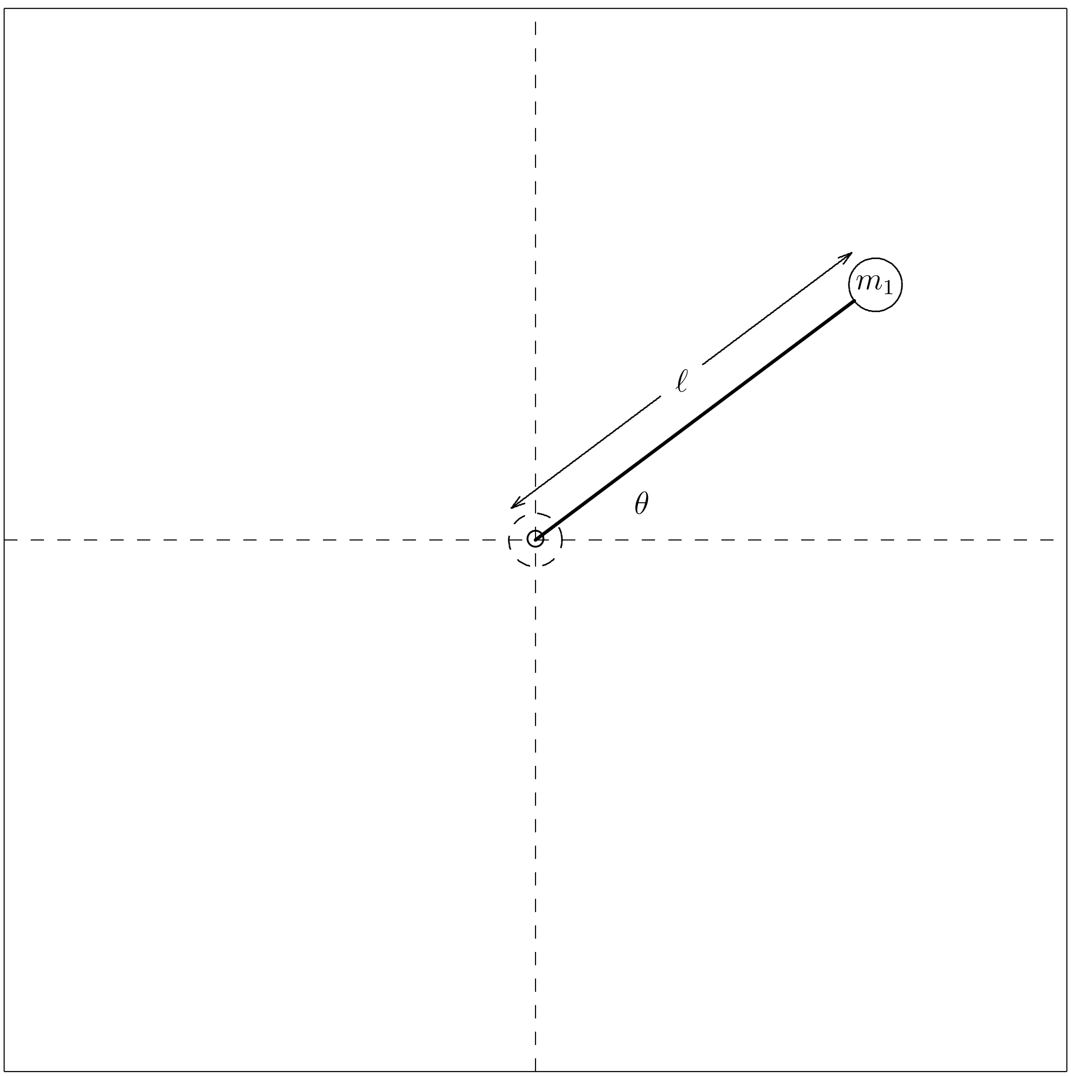
$$\frac{d}{dt}\,\mathcal{L} = \sum_j \ddot{q}_j\,\frac{\partial \mathcal{L}}{\partial \dot{q}_j} + \sum_j \dot{q}_j\,\frac{\partial \mathcal{L}}{\partial q_j} + \frac{\partial \mathcal{L}}{\partial t} \tag{13}$$

This requires that the same sum over degrees of freedom appears in the first term in (12). With F defined as in (12), (7) is still valid.

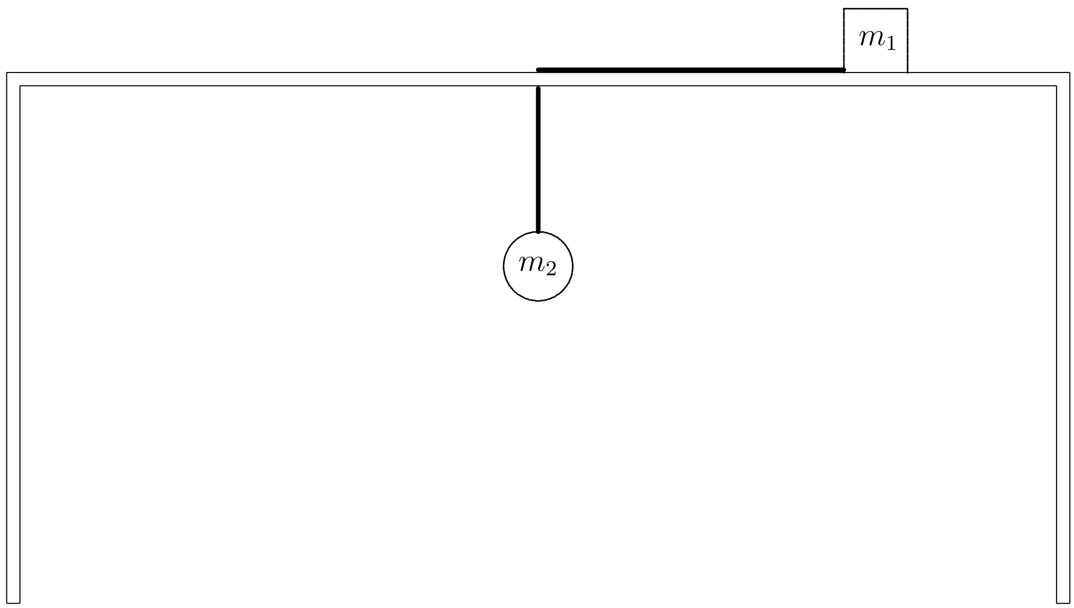
Thus the construction (12) is very general. In fact, it reproduces the usual expression for the energy as the kinetic plus the potential energy whenever the potential depends only on q_j and the kinetic energy is proportional to two powers of the velocity. However, it is really even more general than that. For any Lagrangian that does not depend explicitly on t, (12) defines a conserved quantity. And if the explicit t dependence vanishes because of time translation invariance, the conserved quantity is the energy.

Example - frictionless table

Here is an example of a Lagrangian for a system with two degrees of freedom. Consider a frictionless table in the x-y plane with a hole at the origin. A mass m_1 slides on the surface of the table, but it is attached to a massless string of length R which goes through the hole in the center of the table and hangs straight down where it is attached to a mass m_2. We can describe the configuration of the system by giving the length, ℓ, of string on the table and the angle, θ, of the string on the table from the x axis. From above, this looks like this:

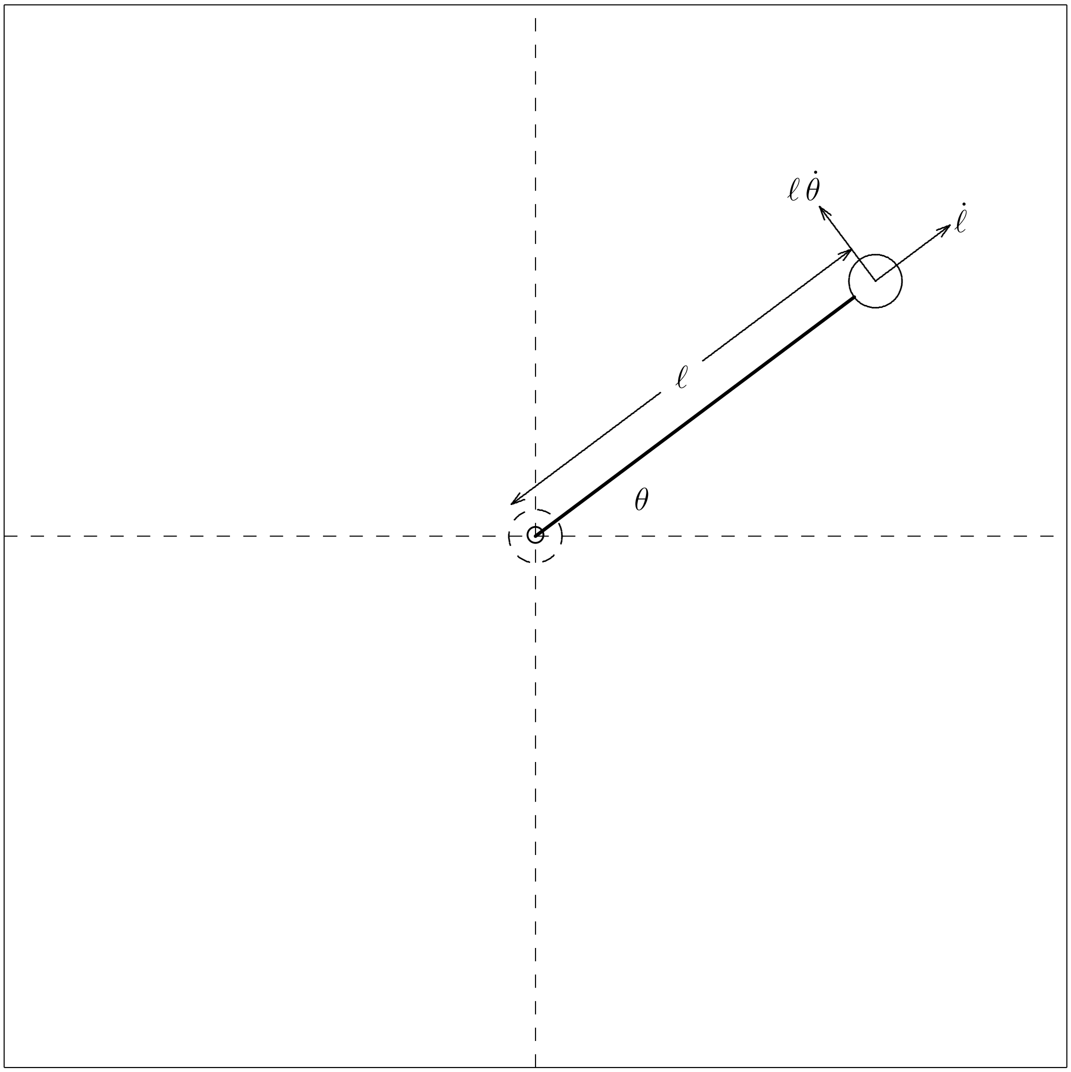


From the side, it looks like



The kinetic energy contains a term from the rate of change of ℓ, proportional to the sum of the masses (because both masses move when ℓ changes), and a term from the rate of change of θ, proportional to m_1 —

$$T(\ell, \theta, \dot{\ell}, \dot{\theta}) = \frac{1}{2}(m_1 + m_2)\dot{\ell}^2 + \frac{1}{2}m_1\ell^2\dot{\theta}^2 \tag{14}$$



There there is also a potential energy related to the height of the mass m_2,

$$U(\ell, \theta, \dot\ell, \dot\theta) = m_2 g \ell \tag{15}$$

Now

$$\mathcal{L}(\ell, \theta, \dot\ell, \dot\theta) = \frac{1}{2}(m_1 + m_2)\dot\ell^2 + \frac{1}{2}m_1 \ell^2 \dot\theta^2 - m_2 g \ell \tag{16}$$

The Euler-Lagrange equations are

$$0 = m_1 \ell \dot\theta^2 - m_2 g - \frac{d}{dt}\Big((m_1 + m_2)\dot\ell\Big) \tag{17}$$

and

$$0 = -\frac{d}{dt}\Big(m_1 \ell^2 \dot\theta\Big) \tag{18}$$

The relation (18) is another example of conservation of a generalized momentum.

$$\mathcal{L}(\ell, \theta, \dot\ell, \dot\theta) = \frac{1}{2}(m_1 + m_2)\dot\ell^2 + \frac{1}{2}m_1 \ell^2 \dot\theta^2 - m_2 g \ell \tag{19}$$

doesn't depend on θ. Thus the generalized momentum corresponding to θ,

$$\frac{\partial \mathcal{L}}{\partial \dot{\theta}} = m_1 \ell^2 \dot{\theta} \tag{20}$$

is constant. This conserved generalized momentum is another example of angular momentum about the origin, which is conserved because the system has a rotation symmetry about the origin.

For this system

$$\mathcal{L}(\ell, \theta, \dot{\ell}, \dot{\theta}) = \frac{1}{2}(m_1 + m_2)\dot{\ell}^2 + \frac{1}{2} m_1 \ell^2 \dot{\theta}^2 - m_2 g \ell \tag{21}$$

there is no explicit time dependence, so we expect a conserved energy. Because there are two degrees of freedom, we have to use the construction (12), which gives for the conventional energy, $T + U$,

$$\begin{aligned} F &= \dot{\ell}\,(m_1 + m_2)\dot{\ell} + \dot{\theta}\, m_1 \ell^2 \dot{\theta} - \frac{1}{2}(m_1 + m_2)\dot{\ell}^2 - \frac{1}{2} m_1 \ell^2 \dot{\theta}^2 + m_2 g \ell \\ &= \frac{1}{2}(m_1 + m_2)\dot{\ell}^2 + \frac{1}{2} m_1 \ell^2 \dot{\theta}^2 + m_2 g \ell \end{aligned} \tag{22}$$

When is F the energy?

To even get started on this, we have to restrict ourselves to situations where there are no frictional forces, and therefore nothing is converting kinetic and potential energy into heat. Since heat is just kinetic energy of random particle motion, this really just means that we need to look at a system at a sufficiently fundamental level to see all the kinetic energy explicitly. For now we will take care of this by restricting ourselves to systems described by Lagrangians of the form $T - V$. Then the first answer is that F is the energy whenever the kinetic energy T is quadratic in the velocities, $\dot{q}$ and the potential energy V does not depend on the velocities. The reason is that the differential operator

$$z \frac{\partial}{\partial z} \tag{23}$$

counts the degree in z, because

$$z \frac{\partial}{\partial z}\, z^n = z\, n\, z^{n-1} = n\, z^n \tag{24}$$

That is, this operator acting on a term that is some power of z just gives the term back multiplied by the power. When we sum over all the velocities, the differential operator

$$\sum_j \dot{q}_j \, \frac{\partial}{\partial \dot{q}_j} \tag{25}$$

adds up the powers of all the velocities. When this acts on a function quadratic in the velocities, every term in the function just gets multiplied by 2. For example

$$\sum_j \dot{q}_j \, \frac{\partial}{\partial \dot{q}_j}\, \dot{q}_1^2 = \dot{q}_1 \, \frac{\partial}{\partial \dot{q}_1}\, \dot{q}_1^2 = 2\dot{q}_1^2 \tag{26}$$

and

$$\sum_j \dot q_j \frac{\partial}{\partial \dot q_j} \dot q_1 \dot q_2 = \dot q_1 \frac{\partial}{\partial \dot q_1} \dot q_1 \dot q_2 + \dot q_2 \frac{\partial}{\partial \dot q_2} \dot q_1 \dot q_2 = \dot q_1 \dot q_2 + \dot q_1 \dot q_2 = 2\dot q_1 \dot q_2 \tag{27}$$

Thus if T is quadratic in the velocities and V is independent of the velocities,

$$\sum_j \dot q_j \frac{\partial \mathcal{L}}{\partial \dot q_j} = \sum_j \dot q_j \frac{\partial}{\partial \dot q_j} (T - V) = 2T \tag{28}$$

and then

$$F = \sum_j \dot q_j \frac{\partial \mathcal{L}}{\partial \dot q_j} - \mathcal{L} = 2T - (T - V) = T + V = E \tag{29}$$

and as promised, F is the energy.

At the fundamental level, the kinetic energy is always quadratic in the velocities, because it is just a sum over all the parts of the system of $\frac{1}{2}mv^2$. But what can go wrong with this can be seen in examples like the bead on a rod rotating with fixed angular velocity ω. Sometimes, we have to separate the variables that describe the system into those that are dynamical - like the position of the bead along the rod - and those that are imposed by some "external" constraint - like the angle of the rod, which is fixed whatever is causing the rod to rotate. Then we include in our list of coordinates only the dynamical coordinates and put the effect of the others into the Lagrangian by hand. Then if our external constraint is causing the motion, some of the velocities that appear in the kinetic energy do not correspond to dynamical coordinates in our Lagrangian, and they are not included in the sum,

$$\sum_j \dot q_j \frac{\partial}{\partial \dot q_j} \tag{30}$$

In such a case, (29) is not correct, and F is not the energy.

In a system with time translation invariance, however, there can be no time dependent external constraint. Such a thing would look different at different times and break time translation invariance. Thus time translation invariance does two things. Not only does it ensure that the Lagrangian does not depend explicitly on time - which implies that F is conserved. But is also ensures that the kinetic energy is quadratic in the dynamical velocities, which ensures that F is the energy. In fact, time translation invariance is the more general answer to the question. Even if $\mathcal{L}$ is not in the form $T - V$, time translation invariance implies that F is the energy.

Symmetries and transformations

What is a symmetry? We have talked about several examples, so perhaps we should define precisely what we mean by it. Symmetry is a mathematical statement of some very specific regularity in a system. A system has a symmetry if there is some transformation you can make that leaves the system looking exactly as it did before. We tend to regard things with many symmetries as pretty, like the kaleidoscope that we saw at the beginning of lecture, which has many planes of symmetry.

In the case of mechanics, we have an even more specific meaning in mind. Let us now consider a class of symmetries in which we make some transformation of the coordinates describing a system at a fixed time. What this means mathematically is that we define a new set of coordinates as functions of the original coordinates. The transformation is then a symmetry if the physics looks exactly the same in terms of the new coordinates as it did in the old coordinates.

We talked briefly about such a transformation when we discussed the double pendulum with two equal masses in lecture 2. The Lagrangian for the double pendulum for small oscillations looks approximately like

$$\frac{m}{2}\left(\dot{x}_1^2+\dot{x}_2^2\right)-\frac{g}{\ell}\left(x_1^2+x_2^2\right)-\frac{K}{2}\left(x_1-x_2\right)^2 \tag{31}$$

This has the property that it is unchanged if we interchange x_1 and x_2. This is the mathematical statement of the obvious physical symmetry of the system.

The symmetry of the double pendulum is an example of a **discrete** symmetry, so-called because the symmetry is an all or nothing sort of thing. The transformation cannot be made bigger or smaller - it is fixed by the structure of the symmetry.

It is more even interesting to consider symmetries in which the symmetry transformation can be made arbitrarily small. Such a thing is called a **continuous** symmetry, because the transformation can change the system continuously. By putting arbitrarily small transformations together, we can get a whole set of transformations which, unlike the symmetry of the double pendulum, depend on a parameter that can be continuously varied.

An example is translations. We think that space probably looks the same everywhere, and we could describe this by saying that there is a symmetry in which we move everything by the same arbitrary vector and we would end up with a completely equivalent physical system.

Here is the general theoretical setup (we'll discuss examples in more detail shortly). Consider a system of n degrees of freedom described by coordinates q_j for $j = 1$ to n. Let's assume that there is a symmetry in which each of the coordinates changes only a tiny bit, proportional to an infinitesimal parameter, ϵ and that the changes involve only the current configuration of the system. What we mean precisely by **infinitesimal** is that ϵ is sufficiently small that we can always ignore terms of order ϵ^2. Translating what we have just assumed into mathematics, we consider a symmetry in which the coordinates q_j are transformed as follows:

$$q_j \to \tilde{q}_j = q_j + \epsilon\,\kappa_{q_j}(q)\,. \tag{32}$$

That is each of the coordinates changes by ϵ times a function $\kappa_{q_j}(q)$ of the qs. The $\kappa_{q_j}(q)$ tells you how the variable q_j changes under the transformation. The transformation (32) is a symmetry of the Lagrangian if

$$\mathcal{L}(\tilde{q},\dot{\tilde{q}}) = \mathcal{L}(q,\dot{q}) \tag{33}$$

Example: space translations for one particle

Here is a simple (perhaps even boring) example. Consider a particle with mass m moving along the x-axis in a potential. When does this system have a symmetry under the infinitesimal transformation

$$x \to \tilde{x} = x + \epsilon\,? \tag{34}$$

This transformation has the form of (32) with $\kappa(x) = 1$. The Lagrangian looks like this:

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2} m \dot{x}^2 - V(x) \tag{35}$$

The condition that (34) is a symmetry is then

$$\mathcal{L}(\tilde{x}, \dot{\tilde{x}}) = \mathcal{L}(x, \dot{x}) \tag{36}$$

Because $\kappa(x)$ is just a constant, we have

$$\dot{\tilde{x}} = \frac{d}{dt} \tilde{x} = \dot{x} \tag{37}$$

and so (36) becomes

$$\mathcal{L}(x + \epsilon, \dot{x}) = \mathcal{L}(x, \dot{x}) \tag{38}$$

for infinitesimal ϵ. Because the $\dot{x}$ doesn't change and kinetic energy is the same on both sides, so this condition only effects V -

$$V(x + \epsilon) = V(x) \tag{39}$$

But because this is supposed to be true for any infinitesimal ϵ, we can use the Taylor expansion (surprise, surprise) to rewrite (39) as

$$V(x + \epsilon) = V(x) + \epsilon V'(x) + \mathcal{O}(\epsilon^2) = V(x) \tag{40}$$

If this is to be satisfied for infinitesimal ϵ, we must have

$$V'(x) = 0 \tag{41}$$

so that $V(x)$ is just a constant and the particle has no force on it at all. In this case, mv is a conserved momentum. We will see how this connection between symmetry and conserved momentum generalizes to more complicated (and more interesting) situations.

Another reason that I wanted to look at this simple system in detail is to emphasize the difference between continuous and discrete symmetries. Suppose that instead of being constant, the potential in (35) is

$$V(x) = -E_0 \cos(x/\ell) \tag{42}$$

where E_0 and ℓ are constants. This system also has a symmetry under space translations of the form

$$x \to \tilde{x} = x + 2\pi n \ell \tag{43}$$

for any integer n. But here we clearly cannot conclude that $V(x)$ is constant because we started with an example with the symmetry that is not constant. Except at special points where the particle is in equilibrium, there is a force on it. There is no conserved momentum (though energy is still conserved because the Lagrangian does not depend explicitly on t). The difference between this and the previous example is that this is a discrete symmetry. The changes in x that leave the Lagrangian invariant are a discrete set. They cannot be varied continuously, and they cannot be made infinitesimally small. Thus we cannot use the Taylor expansion argument to conclude that $V(x)$ is constant.

Space translations for two particles

Space translation symmetry becomes interesting and important when there is more than one particle. Let us now consider a one-dimensional system of 2 particles, with positions x_1 and x_2, so that $q_j = x_j$ for $j = 1$ to 2. A space translation in the x direction just adds the same infinitesimal constant, ϵ, to both x_1 and x_2 — so the transformation has the form

$$x_1 \to \tilde{x}_1 = x_1 + \epsilon \,, \qquad x_2 \to \tilde{x}_2 = x_2 + \epsilon \,. \tag{44}$$

Notice that this satisfies (32), with $q_1 = x_1$, $q_2 = x_2$ and $\kappa_{x_1}(x) = \kappa_{x_2}(x) = 1$. This is a symmetry of any Lagrangian that depends only on $\dot{x}_1$ and $\dot{x}_2$ and the difference between $x_1 - x_2$, for example

$$\mathcal{L}(x, \dot{x}) = \frac{m_1}{2}\,\dot{x}_1^2 + \frac{m_2}{2}\,\dot{x}_2^2 - V(x_1 - x_2) \tag{45}$$

The transformation (44) is a symmetry of the kinetic energy because ϵ is a constant, so that

$$\dot{\tilde{x}}_1 = \dot{x}_1 \,, \qquad \dot{\tilde{x}}_2 = \dot{x}_2 \,. \tag{46}$$

It is a symmetry of the potential energy because the ϵs cancel when we subtract one coordinate from another, so that

$$\tilde{x}_1 - \tilde{x}_2 = (x_1 + \epsilon) - (x_2 + \epsilon) = x_1 - x_2 \,. \tag{47}$$

Putting (46) and (47) together implies

$$\mathcal{L}(\tilde{x}, \dot{\tilde{x}}) = \mathcal{L}(x, \dot{x}) \,. \tag{48}$$

This system has a conserved momentum,

$$p = m_1\dot{x}_1 + m_2\dot{x}_2 \tag{49}$$

because the forces that come from the potential energy obey Newton's third law. This in turn is related to the fact that the potential energy depends only on $x_1 - x_2$, which in turn is related to the symmetry. We will see that this connection between a continuous symmetry and the existence of a conserved momentum is a general thing.

Space translations for many particles

Let us now consider a one-dimensional system of n particles, with positions x_j, so that $q_j = x_j$ for $j = 1$ to n. A space translation in the x direction adds the same infinitesimal constant, ϵ, to each x_j — so the transformation has the form

$$x_j \to \tilde{x}_j = x_j + \epsilon \;\forall n \,. \tag{50}$$

Notice that this satisfies (32), with $q_j = x_j$ and all of the $\kappa_{x_j}(x) = 1$. This is a symmetry of any Lagrangian that depends only on $\dot{x}_j$ and differences between two x_js, for example

$$\mathcal{L}(x, \dot{x}) = \sum_j \frac{m_j}{2}\,\dot{x}_j^2 - V(x_1 - x_2\,,\, x_2 - x_3\,,\, \cdots\,,\, x_j - x_{j+1}\,,\, \cdots\,,\, x_{n-1} - x_n) \,. \tag{51}$$

Again, (50) is a symmetry of the kinetic energy because ϵ is a constant, so that

$$\dot{\tilde{x}}_j = \dot{x}_j \, . \tag{52}$$

It is a symmetry of the potential energy because the ϵs cancel when we subtract one coordinate from another, so that

$$\tilde{x}_j - \tilde{x}_{j+1} = (x_j + \epsilon) - (x_{j+1} + \epsilon) = x_j - x_{j+1} \, . \tag{53}$$

Here the total momentum,

$$\sum_j m_j v_j \tag{54}$$

is conserved because the forces that come from the potential energy obey Newton's third law.

Finding Symmetries

So far, we have looked at systems in which it is pretty obvious what the symmetry transformation is. But when come upon some Lagrangian, you may want to find **what transformation the Lagrangian is invariant under.** The way to do this is to write down how your Lagrangian transforms under a general infinitesimal transformation of the form

$$q_j \to \tilde{q}_j = q_j + \epsilon \, \kappa_{q_j}(q) \, . \tag{55}$$

and then require that it be invariant - that is that the coefficient of the ϵ term in

$$\mathcal{L}(\tilde{q}, \dot{\tilde{q}}) - \mathcal{L}(q, \dot{q}) \tag{56}$$

vanishes. Here is a simple example, where you could probably guess the answer, but it will illustrate the technique.

$$\mathcal{L}(x, \dot{x}) = \frac{m_1}{2} \, \dot{x}_1^2 + \frac{m_2}{2} \, \dot{x}_2^2 - V(x_1 + 2x_2) \tag{57}$$

Does this Lagrangian has a symmetry? To see, we want to find $\mathcal{L}(\tilde{x}, \dot{\tilde{x}})$ where

$$x_j \to \tilde{x}_j = x_j + \epsilon \, \kappa_{x_j}(x) \, . \tag{58}$$

$$\mathcal{L}(\tilde{x}, \dot{\tilde{x}}) = \frac{m_1}{2} \left(\dot{x}_1 + \epsilon \dot{\kappa}_{x_1} \right)^2 + \frac{m_2}{2} \left(\dot{x}_2 + \epsilon \dot{\kappa}_{x_2} \right)^2 - V(x_1 + \epsilon \kappa_{x_1} + 2x_2 + 2\epsilon \kappa_{x_2}) \tag{59}$$

$$= \frac{m_1}{2} \left(\dot{x}_1^2 + 2\epsilon x_1 \dot{\kappa}_{x_1} + \cdots \right) + \frac{m_2}{2} \left(\dot{x}_2^2 + 2\epsilon x_2 \dot{\kappa}_{x_2} + \cdots \right) - V\Big(x_1 + 2x_2 + \epsilon(\kappa_{x_1} + 2\kappa_{x_2}) \Big) \tag{60}$$

The ϵ term will vanish if

$$\dot{\kappa}_{x_1} = \dot{\kappa}_{x_2} = \kappa_{x_1} + 2\kappa_{x_2} = 0 \tag{61}$$

One way to satisfy the first two conditions is to take the κs to be constant independent of x (if they depend on x, the time derivatives would be nonzero). Then the last condition implies that we can take

$$\kappa_{x_1} = -2 \qquad \kappa_{x_2} = 1 \tag{62}$$

which gives a transformation of the form

$$x_1 \to x_1 - 2\epsilon \qquad x_2 \to x_2 + \epsilon \tag{63}$$

You see that it is sometimes easy to find symmetry transformations. It is harder to show that you have found them all. This depends on details, in this case the precise form of the function V. We won't talk about that now.

Rotations

Another simple example can be found in the example we discussed earlier of the mass m_1 sliding on a horizontal frictionless table, connected to a string that goes through a hole at the origin to a mass m_2 hanging below the table with Lagrangian (16). This physical system has a symmetry under rotations about the z axis, which add a constant to the angle θ without changing ℓ,

$$\ell \to \tilde{\ell} = \ell\,, \qquad \theta \to \tilde{\theta} = \theta + \epsilon\,. \tag{64}$$

This symmetry is responsible for the fact that the Lagrangian does not depend on θ at all, but only on $\dot{\theta}$, because the condition for symmetry, that

$$\mathcal{L}(\tilde{\ell}, \tilde{\theta}, \dot{\tilde{\ell}}, \dot{\tilde{\theta}}) = \mathcal{L}(\ell, \theta, \dot{\ell}, \dot{\theta}) \tag{65}$$

in this case becomes

$$\mathcal{L}(\ell, \theta + \epsilon, \dot{\ell}, \dot{\theta}) = \mathcal{L}(\ell, \theta, \dot{\ell}, \dot{\theta}) \tag{66}$$

This just says that if we make a little change in θ in the function, nothing happens, so the function must not depend on θ. If we wanted to say the same thing in fancier mathematics, we could say that because ϵ is infinitesimal, we can reliably Taylor expand the left-hand-side of (66) and keep only the first two terms,

$$\mathcal{L}(\ell, \theta + \epsilon, \dot{\ell}, \dot{\theta}) = \mathcal{L}(\ell, \theta, \dot{\ell}, \dot{\theta}) + \epsilon\,\frac{\partial}{\partial\theta}\mathcal{L}(\ell, \theta, \dot{\ell}, \dot{\theta}) \tag{67}$$

Putting (67) into (66) implies

$$\frac{\partial}{\partial\theta}\mathcal{L}(\ell, \theta, \dot{\ell}, \dot{\theta}) = 0 \tag{68}$$

which as promised is the statement that $\mathcal{L}$ does not depend on θ.

As we saw in lecture 7, this system also has a conserved quantity, the angular momentum

$$L = m_1 \ell^2 \dot{\theta}\,. \tag{69}$$

So again, we see a connection between a symmetry and a conserved quantity.

Noether's theorem

Having seen a correlation between symmetry and conservation laws in a couple of examples, let us now consider how this works in more generality and see if we can pin down the connection precisely. Putting (32) into (33) and Taylor expanding gives

$$\mathcal{L}(q,\dot{q}) = \mathcal{L}(q+\epsilon\kappa,\dot{q}+\epsilon\dot{\kappa}) = \mathcal{L}(q,\dot{q}) + \epsilon\sum_{q_j}\dot{\kappa}_{q_j}\frac{\partial\mathcal{L}}{\partial\dot{q}_j} + \epsilon\sum_{q_j}\kappa_{q_j}\frac{\partial\mathcal{L}}{\partial q_j}\,. \tag{70}$$

Thus the condition that (32) is a symmetry of the Lagrangian is equivalent to

$$\sum_{q_j}\dot{\kappa}_{q_j}\frac{\partial\mathcal{L}}{\partial\dot{q}_j} + \sum_{q_j}\kappa_{q_j}\frac{\partial\mathcal{L}}{\partial q_j} = 0 \tag{71}$$

For example, for space translations, (50),

$$\kappa_{x_j} = 1 \tag{72}$$

for all j, because all the x_js are translated in the same way. Thus $\dot{\kappa} = 0$ and the condition of symmetry becomes

$$\sum_{x_j}\kappa_{x_j}\frac{\partial\mathcal{L}}{\partial x_j} = \sum_{x_j}\frac{\partial\mathcal{L}}{\partial x_j} = 0 \tag{73}$$

This is equivalent to the statement that $\mathcal{L}$ depends only on differences $x_j - x_k$.

Now return to the general case and consider the quantity

$$\sum_{q_j}\kappa_{q_j}\frac{\partial\mathcal{L}}{\partial\dot{q}_j} \tag{74}$$

Consider the time derivative of this quantity

$$\frac{d}{dt}\sum_{q_j}\left(\kappa_{q_j}\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right) = \sum_{q_j}\dot{\kappa}_{q_j}\frac{\partial\mathcal{L}}{\partial\dot{q}_j} + \sum_{q_j}\kappa_{q_j}\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_j} \tag{75}$$

If we apply the Lagrange equations of motion to the second term, it becomes

$$\frac{d}{dt}\sum_{q_j}\left(\kappa_{q_j}\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right) = \sum_{q_j}\dot{\kappa}_{q_j}\frac{\partial\mathcal{L}}{\partial\dot{q}_j} + \sum_{q_j}\kappa_{q_j}\frac{\partial\mathcal{L}}{\partial q_j} = 0 \tag{76}$$

which vanishes because of the condition (71) that the Lagrangian is symmetric. This is a very important general theorem. It is the precise connection between a continuous symmetry and a conservation law. For every continuous symmetry of the form (32), there is a conservation law — the quantity of the form (74) is constant.

Now we can find the conserved quantity for any symmetry. For example for the Lagrangian of (57), where the κs are given by (62), the conserved quantity is

$$\kappa_{x_1}\frac{\partial\mathcal{L}}{\partial\dot{x}_1} + \kappa_{x_2}\frac{\partial\mathcal{L}}{\partial\dot{x}_2} = -2m_1\dot{x}_1 + m_2\dot{x}_2 \tag{77}$$

This theorem is a special favorite of mine because a lot of my own work in particle physics is based on it. It was worked out early in this century by the great woman mathematician and theoretical physicist Emmy Noether.[1]

The quantity

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \tag{78}$$

that appears in Noether's theorem is a very important one in classical mechanics. It is the "generalized momentum" we talked about earlier, corresponding to the coordinate q_j, we discussed in the last lecture. Of course, if q_j is a normal space coordinate (like x_j), and the kinetic energy has the standard form, it is the momentum (or a component of it). In fact, one way of thinking about (76) is to recognize that in a sense (which I will not explain in detail) (74) is the generalized momentum associated with, ϵ, the infinitesimal variable that describes how all the q_js change under the symmetry.

Momentum conservation from Noether's theorem

For space translation symmetry, because all the κ_{x_j} are equal to 1, the quantity (74) becomes

$$\sum_{x_j} \kappa_{x_j} \frac{\partial \mathcal{L}}{\partial \dot{x}_j} = \sum_{x_j} \frac{\partial \mathcal{L}}{\partial \dot{x}_j} \tag{79}$$

which is just the sum over the momenta of all the individual particles. **Space translation invariance implies that the total momentum is conserved.**

This analysis can obviously be extended to three dimensions, where the coordinates and the corresponding momenta become vectors.

More on rotations

Here is a more involved example that we can discuss if we have time. When we thought about rotations for the bead sliding on the table, we were already using polar coordinates. This makes rotations easy, because an infinitesimal rotation is just a translation of θ,

$$\theta \to \theta' = \theta + \epsilon\,. \tag{80}$$

This is an example with two particles in Cartesian coordinates. Consider two particles, with masses m_1 and m_2, which move in a plane with coordinates $\vec{r}_1 = (x_1, y_1)$ and $\vec{r}_2 = (x_2, y_2)$. Consider the following Lagrangian:

$$\frac{1}{2} m_1 \left(\dot{x}_1^2 + \dot{y}_1^2 \right) + \frac{1}{2} m_2 \left(\dot{x}_2^2 + \dot{y}_2^2 \right) - V\left(x_1^2 + y_1^2 + x_2^2 + y_2^2 - 2(x_1 x_2 + y_1 y_2) \right) \tag{81}$$

[1]For more information about Noether, see **The Life and Times of Emmy Noether** by Nina Byers — http://xxx.lanl.gov/abs/hep-th/9411110 and on the 16 website in handouts/noether.pdf.

This Lagrangian is invariant under the following infinitesimal transformation:

$$\begin{aligned} x_1 \rightarrow \tilde{x}_1 &= x_1 - \epsilon y_1 \qquad & y_1 \rightarrow \tilde{y}_1 &= y_1 + \epsilon x_1 \\ x_2 \rightarrow \tilde{x}_2 &= x_2 - \epsilon y_2 \qquad & y_2 \rightarrow \tilde{y}_2 &= y_2 + \epsilon x_2 \end{aligned} \tag{82}$$

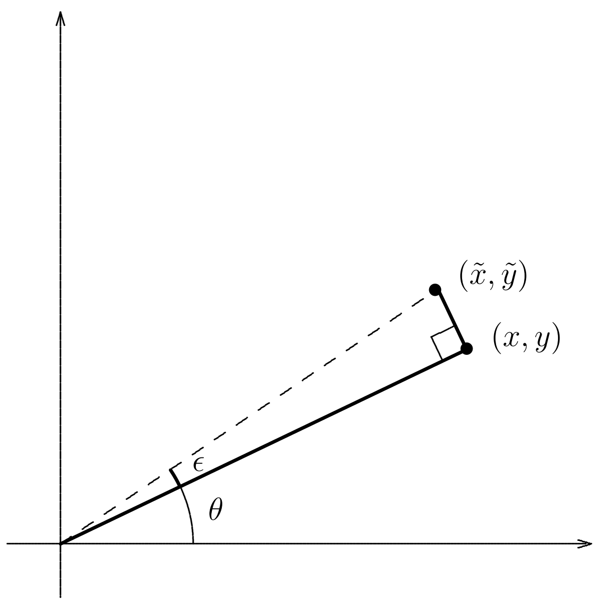
If you have never seen anything like this before, it is not obvious at all, but you can see it by explicit calculation. For example, look at the m_1 terms in the kinetic energy. First note that

$$x_1 \rightarrow \tilde{x}_1 = x_1 - \epsilon y_1 \qquad y_1 \rightarrow \tilde{y}_1 = y_1 + \epsilon x_1 \tag{83}$$

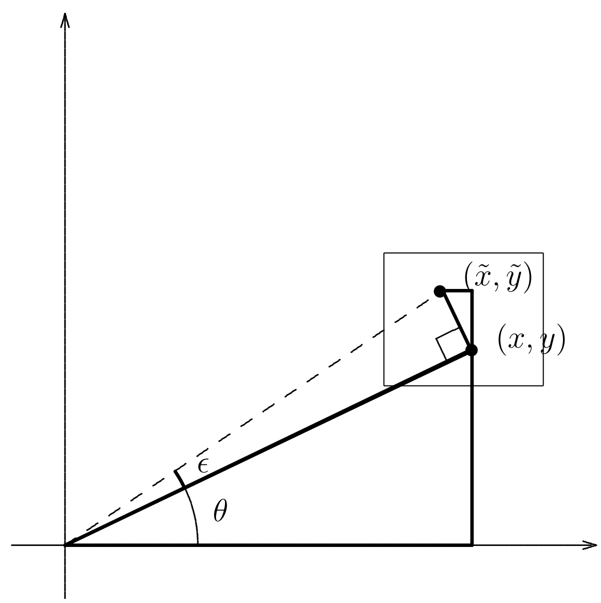
then substitute

$$\dot{\tilde{x}}_1^{\,2} + \dot{\tilde{y}}_1^{\,2} = \Big(\dot{x}_1 - \epsilon \dot{y}_1\Big)^2 + \Big(\dot{y}_1 + \epsilon \dot{x}_1\Big)^2 = \dot{x}_1^2 + \dot{y}_1^2 + \mathcal{O}(\epsilon^2) \tag{84}$$

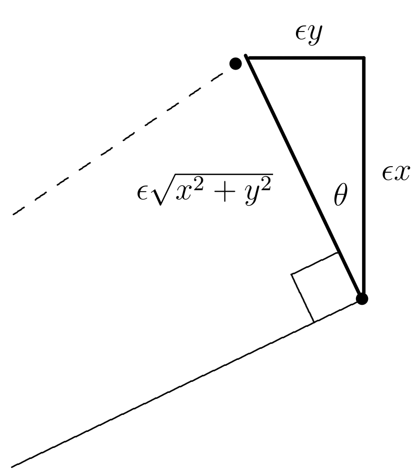
The ϵ^2 terms do not cancel, but we don't care about them because ϵ is infinitesimal - we only consider linear terms in Noether's theorem. So this is good enough. All the other terms work similarly. You can see what the rotation looks like pictorially in the figure below.



The heavy lines represent the vector $\vec{r} = (x, y)$ from the origin to the original point, and the smaller vector from $\vec{r}$ to $\vec{\tilde{r}} = (\tilde{x}, \tilde{y})$. Because ϵ is infinitesimal, these two are nearly perpendicular, and the ratio of their lengths is about ϵ. It is helpful to complete these two lines into a similar triangle.



Then blowing up the box, you can see that the change in x under the rotation is $-\epsilon y$ and the change in y is ϵx.



Now let's see what Noether's theorem looks like. The κs are

$$\kappa_{x_1} = -y_1\,, \quad \kappa_{y_1} = x_1\,, \quad \kappa_{x_2} = -y_2\,, \quad \kappa_{y_2} = x_2\,. \tag{85}$$

and the momenta are

$$\frac{\partial \mathcal{L}}{\partial \dot{x}_j} = m_j \dot{x}_j\,, \quad \frac{\partial \mathcal{L}}{\partial \dot{y}_j} = m_j \dot{y}_j\,. \tag{86}$$

Thus the conserved quantity is

$$\sum_j m_j(x_j\dot{y}_j - y_j\dot{x}_j) \tag{87}$$

As we will see in more detail later in the course, this is the angular momentum.

Notice also that the complicated function of the variables that appears in V can be written as

$$(\vec{r}_1 - \vec{r}_2\,)\cdot(\vec{r}_1 - \vec{r}_2\,) = |\vec{r}_1 - \vec{r}_2\,|^2 \tag{88}$$

just the square of the length of the difference $\vec{r}_1 - \vec{r}_2$.

What functions are invariant?

We have seen some examples of functions that are invariant under transformations, such as the potential in (51),

$$V(x_1{-}x_2\,,\,x_2{-}x_3\,,\,\cdots\,,\,x_j{-}x_{j+1}\,,\,\cdots\,,\,x_{n-1}{-}x_n)\,. \tag{89}$$

It always straightforward to check that a given function is invariant under a specific transformation. And we have seen how to find transformations that are symmetries, if they exist. It is also important to be able to go the other way and to construct the most general function invariant under a transformation or set of transformations.

To illustrate what I am talking about, let's show that a function of $F(x_1, x_2)$ that is invariant under the symmetry transformation

$$x_j \to x_j + \epsilon \tag{90}$$

for $j = 1$ and 2 actually only depends on the difference, $x_1 - x_2$. One way to do this is to change variables to include the variable $x_1 - x_2$, and eliminate x_1 —

$$y \equiv x_1 - x_2 \qquad\qquad x_1 = x_2 + y \tag{91}$$

Then we can define a new function

$$G(y, x_2) \equiv F(x_2 + y, x_2) \tag{92}$$

in terms of the original function. But now, in terms of the new variables the transformation (90) is

$$y \to y\,, \qquad x_2 \to x_2 + \epsilon\,. \tag{93}$$

Now since F is invariant, G must be also, because we have only relabeled things. Thus

$$G(y, x_2 + \epsilon) = G(y, x_2) \tag{94}$$

for infinitesimal ϵ. In words, this says that making an infinitesimal change in x_2 doesn't affect the function, so it is probably obvious that this means that $G(y, x_2)$ is independent of x_2. But if we want to be more formal about it, we can use the Taylor expansion to get

$$G(y, x_2 + \epsilon) = G(y, x_2) + \epsilon\frac{\partial}{\partial x_2}G(y, x_2) + \cdots = G(y, x_2) \tag{95}$$

and thus

$$\frac{\partial}{\partial x_2} G(y, x_2) = 0 \tag{96}$$

which means that G doesn't depend on x_2, so we can take x_2 to be anything in G. Thus using (91)

$$F(x_1, x_2) = G(x_1 - x_2, x_2) = G(x_1 - x_2, 0) = f(x_1 - x_2) \,. \tag{97}$$

It is easy to extend this proof to show that the most general function invariant under the transformation (90) for n variables, x_j for $j = 1$ to n is given by (89).

Notice that what we are doing here makes Noether's theorem seem a little trivial. If all we do with invariance is to show that there is some variable that the function doesn't depend on, then we could have changed variables first and then found the conserved quantity by just using the statement that the generalized momentum associated with a variable that doesn't appear in the Lagrangian is conserved. And in fact, in this course, you can always do that. But there is actually more to Noether's theorem, because transformations that leave a system invariant have an additional interesting property. They form what mathematicians form a group. The group property is quite powerful and often allows you to extend the infinitesimal transformations that we start with to a much larger set. For example, in the case of translations, group theory can be used to show that if things are invariant under (90) for infinitesimal ϵ, they are also invariant for finite ϵ. Then we don't need the Taylor expansion any more. For example, in (94), we could first set x_2 to zero to get

$$G(y, \epsilon) = G(y, 0) \tag{98}$$

and then relabel $\epsilon \to x_2$ to get

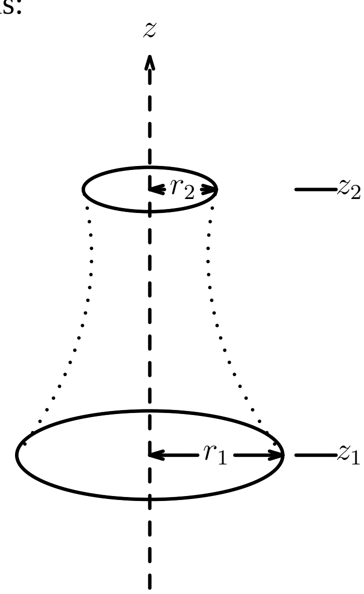
$$G(y, x_2) = G(y, 0) \tag{99}$$

which directly gives (97). We won't pursue this approach much in the course but I may not be able to restrain myself from talking more about it, because it forms the basis of much of the research that I have done in my scientific career.

Example of functionals - Soap bubbles

I thought that it would be interesting to discuss further the pretty mathematics of calculus of variations in a slightly different context. I hope this may give you a better feel for it. There are many examples of the use of the calculus of variations. One nice one is to the shape of soap bubbles. The connection here is that soap bubbles have a surface tension, so they want to minimize their surface area. But the surface area is a functional of the function that describes the shape of the bubble. So this is a job for the calculus of variations. Here's a specific problem that is kind of fun. Consider a bubble formed between two circular loops of wire, both centered on the z axis, one in the $z = z_1$ plane with radius r_1, and the other in the $z = z_2$ plane with radius r_2. We'll assume that $z_1 < z_2$.

So it looks something like this:



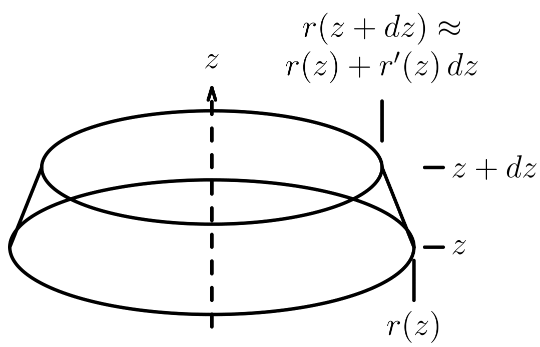
(100)

The soap bubble will be some surface of revolution, because of the cylindrical symmetry of the system , indicated by the dotted line in (100). Thus we can specify the shape of the bubble by giving the radius as a function of z, $r(z)$, subject to the constraint

$$r(z_1) = r_1 \qquad r(z_2) = r_2 \tag{101}$$

which is just the physical requirement that the bubble is attached to the frame — but it looks just like the condition we impose on trajectories in Hamilton's principle. Thus we can compute the shape exactly by solving an Euler-Lagrange equation.

We want to minimize the area of the surface of revolution described by $r(z)$ from $z = z_1$ to $z = z_2$ subject to the constraint (101). First we must compute the area. Consider the area of a small band of surface between z and $z + dz$. The circumference of the band is approximately $2\pi r(z)$ (the difference between $r(z)$ and $r(z + dz)$ does not matter here). The width of the band can be computed by looking at a slice through the z axis, which looks like this:



(102)

Evidently, the width of the slice is approximately

$$\sqrt{dz^2 + dr^2} = dz\,\sqrt{1 + r'(z)^2} \tag{103}$$

Thus the area of the slice is

$$dz\,\alpha\Big(r(z),r'(z)\Big) \quad \text{where} \quad \alpha(r,r') \equiv 2\pi r\,\sqrt{1+r'^2} \tag{104}$$

and the total area is

$$A[r] = \int_{z_1}^{z_2} \alpha\Big(r(z),r'(z)\Big)\,dz = 2\pi\int_{z_1}^{z_2} r(z)\,\sqrt{1+r'(z)^2}\,dz \tag{105}$$

Now we can use the calculus of variations. The mathematics of this problem is the same as in finding the vanishing variation the action, $S[x]$, so the condition is the analog of the Lagrange equation,

$$\frac{\delta A[r]}{\delta r(z)} = 0 \tag{106}$$

We can do the functional differentiations in the same way as above to the analog of the Lagrange equation —

$$\frac{\partial}{\partial r}\alpha(r,r') - \frac{d}{dz}\frac{\partial}{\partial r'}\alpha(r,r') = 0 \tag{107}$$

But rather than trying to solve the Euler-Lagrange equation directly, it is useful to extend the analogy with Lagrangian mechanics and notice that because α does not depend explicitly on z, there will be an analog of energy that is independent of z:

$$\dot q\,\frac{\partial\mathcal{L}}{\partial\dot q} - \mathcal{L} \to r'\,\frac{\partial\alpha}{\partial r'} - \alpha \tag{108}$$

$$r'\,\frac{\partial\alpha}{\partial r'} - \alpha = r'\,\frac{\partial}{\partial r'}\left(2\pi r\sqrt{1+r'^2}\right) - 2\pi r\sqrt{1+r'^2} \tag{109}$$

$$= r'\,2\pi r\,\frac{r'}{\sqrt{1+r'^2}} - 2\pi r\sqrt{1+r'^2} \tag{110}$$

$$= 2\pi r\,\left(\frac{r'^2}{\sqrt{1+r'^2}} - \frac{1+r'^2}{\sqrt{1+r'^2}}\right) = -\frac{2\pi r}{\sqrt{1+r'^2}} \tag{111}$$

$$\frac{r}{\sqrt{1+r'^2}} = r_0 \quad \text{for some constant } r_0 \tag{112}$$

We could mess around with this — but it easier to guess the answer.

$$r(z) = r_0\cosh\tfrac{z-b}{r_0} \tag{113}$$

$$r'(z) = r_0\,\frac{1}{r_0}\sinh\tfrac{z-b}{r_0} = \sinh\tfrac{z-b}{r_0} \tag{114}$$

Now the key step that might actually cause you to make this guess is the next one.

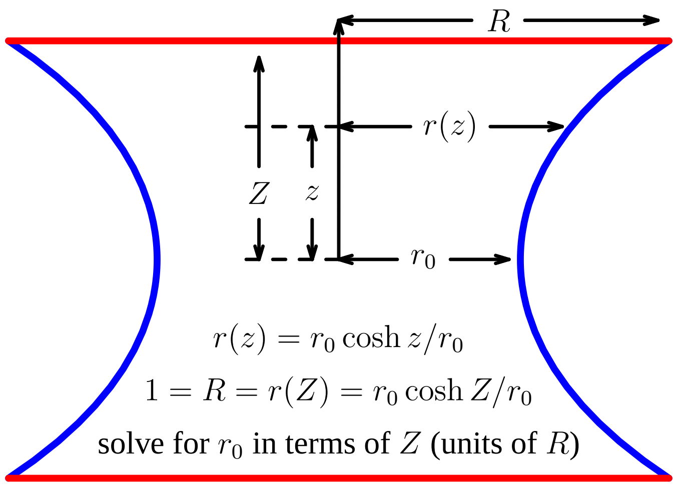
$$1+r'(z)^2 = 1+\sinh^2\tfrac{z-b}{r_0} = \cosh^2\tfrac{z-b}{r_0} \tag{115}$$

$$\frac{r}{\sqrt{1+r'^2}} = \frac{r_0\cosh\frac{z-b}{r_0}}{\cosh\frac{z-b}{r_0}} = r_0 \tag{116}$$

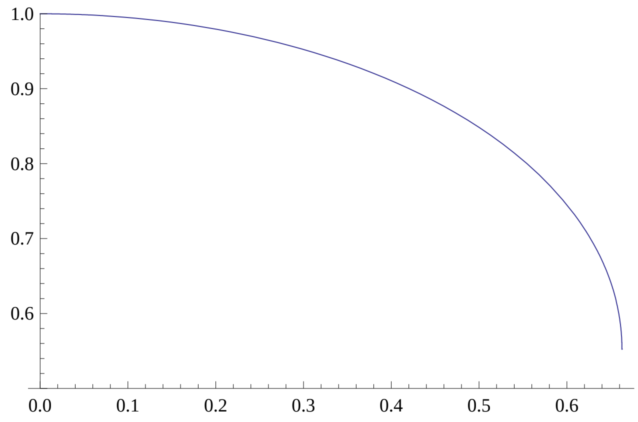
In the demo I will show you, both rings have the same radius, $r_1 = r_2 = R$. Then we define the z coordinate system so that $z = 0$ is right in the middle. Then the fact that $\cosh(x) = \cosh(-x)$ implies that $b = 0$ so that

$$r(z) = r_0 \cosh(z/r_0) \tag{117}$$

and the situation looks like the diagram below (the red lines represent a side view of the rings and the blue curves are a slice through the bubble):

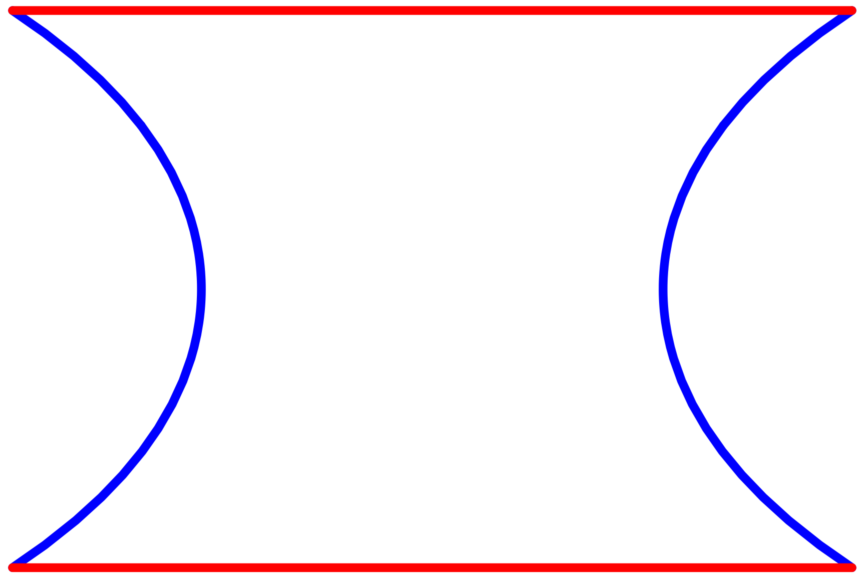


One can analyze this analytically, but it is complicated, and since we have *Mathematica* around, we may as well just find an numerical solution for r_0 as a function of the ring separation Z for fixed ring radius R. This is shown below and in the notebook **soapbubble.nb**.



(118)

Note that something happens if you try to make Z/R big. The maximum is $\approx 0.6627\, R$, for which the bubble looks something like the blue curves:



For large Z/R (as you can see analytically in the appendix), there is no real solution for R/r_0, which means that we cannot actually minimize the area in this way. Physically, what is going on is that for large z it is always energetically favorable to just keep narrowing in the center until the bubble actually breaks into two separate bubbles on the two rings. We should be able to see this in the demo. You can also get a sense of what is happening by looking at **soapbubble.nb**, increasing Z/R to its maximum value, and then clicking on color 0, which toggles diagonal lines that represent a cross-section of the cone through the two circles. The system is unstable when the bubble gets beyond that cone.

A Analytic treatment of the bubble

Now

$$R/r_0 = \cosh(Z/r_0) = \cosh\Big((R/r_0)(Z/R)\Big) \tag{119}$$

is a transcendental equation that can be solved numerically for aR in terms of Z/R. Because, from (117),

$$r(0) = r_0 \tag{120}$$

the physical interpretation of r_0 is that it is the radius of the bubble at its narrowest point, in the center. When we are finished messing with the math, we will express things in term of r_0/R, because that is easier to think about physically.

It is actually much easier to solve for Z/R in terms of R/r_0, because (119) can be rewritten as a quadratic equation for $e^{(R/r_0)(Z/R)}$ —

$$\begin{aligned} R/r_0 = \cosh((R/r_0)(Z/R)) &= \left(e^{(R/r_0)(Z/R))} + e^{-(R/r_0)(Z/R))}\right)/2 \\ \Rightarrow e^{2(R/r_0)(Z/R)} - 2R/r_0\, e^{(R/r_0)(Z/R)} + 1 &= 0 \end{aligned} \tag{121}$$

with solution

$$e^{(R/r_0)(Z/R)} = R/r_0 + \sqrt{(R/r_0)^2 - 1} \tag{122}$$

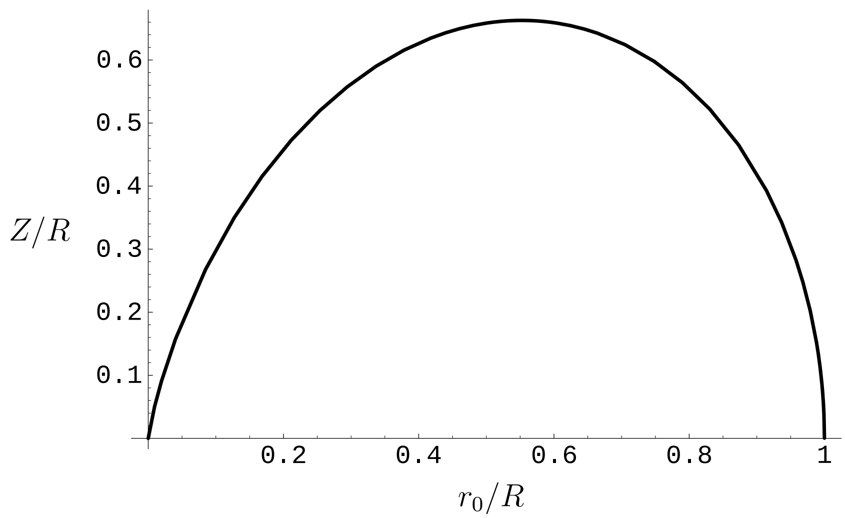
or

$$Z/R = \frac{\ln\left(R/r_0 + \sqrt{(R/r_0)^2 - 1}\right)}{R/r_0} \tag{123}$$

Expressing this in terms of r_0 we have

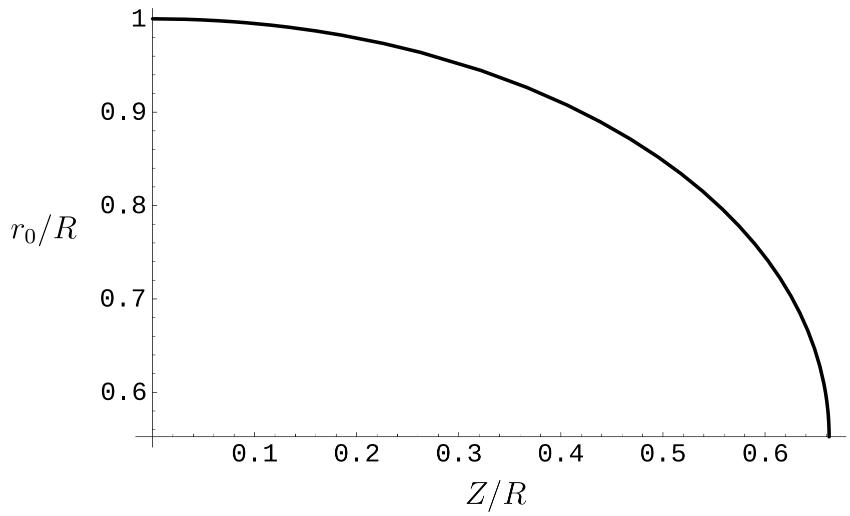
$$\frac{Z}{R} = \frac{r_0}{R}\ln\left(R/r_0 + \sqrt{(R/r_0)^2 - 1}\right) \tag{124}$$

This shows a plot of Z/R versus r_0/R



The first thing to notice about thus graph is that there are two possible values of r_0/R for each values of Z/R. Physically, this doesn't make sense. We can understand what is going by looking at the ends of the r_0/R axis on the first graph. The rightmost point on the graph corresponds to $Z = 0$ and $r_0 = R$ which is the right solution because in the limit of very small Z the area is obviously minimized by an almost flat ribbon between to two rings. Conversely, the leftmost point, $z = 0$ and $r_0 = 0$ is obviously not a physical solution. While the variation vanishes there, the area is not minimized. The leftmost point is some kind of saddle point. Now as we go up along the Z/R axis, the same situation obtains. The point on the right corresponds to the minimum. Thus the physical region of the first graph is only the part to the right of the maximum, which occurs for $r_0 = r_{\min} \approx 0.5524\, R$.

Now that we know what is happening, we can display just the physical solution by plotting r_0/R versus Z/R from $r_{\rm min}/R$ to 1.



This is the analytic version of (118).

lecture 10

Topics:

The structure of science and common sense
The speed of light
Time dilation
The twin paradox
The Doppler effect
The twin paradox and inertial frames

The structure of science and common sense

from www.raremaps.com

The beautiful edifice of Newtonian mechanics, which we have seen a bit of in the last few weeks, provides a wonderful precise mathematical description of most of the things we see in our everyday world. It is obviously right. But it is also wrong. We have discussed qualitatively the underlying quantum mechanical reality from which Newton's mechanics emerges as an approximation. In the next few weeks, we will discuss in quantitative detail the bizarre things that go wrong with Newton's picture at large velocities. What is going on here? Newtonian mechanics beautifully captures the mathematical essence of what we know about the world in which we have grown up. Once we get used to the mathematical language, it is perfectly in accord with our common sense understanding of the world. We feel in our bones that it is right. How can it be wrong?

But it is wrong. I am going to tell you today that it is wrong and what is right and you will not understand me or believe it. Even if you have heard this before, and you think that you have internalized it, you are still not going to really understand it or believe it. In fact, you will not even have any sense of what it would mean to understand it or believe it. It is that strange. It doesn't make any sense.

The first thing to say is that there is no reason why our sense should have anything at all to do with what happens at extreme conditions, far from what we are used to in everyday life. We have some direct intuition about things that are about our size – and maybe a few powers of ten bigger and smaller. We can feel in our bones what happens for accelerations not much different than g, and velocities like those we are used to. But if we go far outside this familiar range of parameters, it would be rather surprising if our common sense worked very well. We should be prepared for surprises. If anything, what should surprise us is that our common sense works as far as it does. We have to go to really enormous velocities, on our everyday scale, before Newtonian kinematics starts to break down. And atoms, which exhibit quantum behavior in all its glory, are very small. This is a theme that we will return to several times.

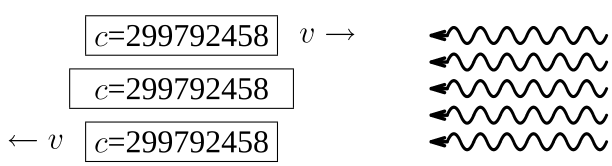
The wonderful thing about the discipline of modern science is that we can say sensible things about phenomena even when our sense doesn't work. We do this by keeping ourselves firmly grounded in what we understand, but at the same time recognizing the limitations of our knowledge. It is useful to think of science as a map of a peculiar space - the space of parameters that describe physical phenomena, things like size, mass, speeds, temperatures, etc. We have all grown up in a familiar, comfortable neighborhood described by a small region on this map. But we have expanded our knowledge of the terrain in much the same way that ancient explorers improved their maps of the known world. We work our way out from what we know into the unknown, exploring and pushing the boundary of what we know farther and farther in different directions away from the range of phenomena that we see in the everyday world. In the next few weeks, we are going to discuss one of these directions — the realm of the very fast. When it is strange, don't be surprised. The reason that you don't understand is very simple. It is that you are slow! Not mentally slow, but physically slow. You have spent all your life moving at speeds very very tiny compared to the speed of light, so nothing in your experience has prepared you for the phenomena that happen all the time at large speeds (which we call "relativistic" speeds — a rather bizarre grammatical construction if you think about it, but standard). Try to bear that in mind when it seems that what we are doing doesn't make any sense.

The speed of light

The speed of light is exactly 299,792,458 m/s. What "exactly" means in this case is just what it says. Because the speed of light, as we will see, is is a kind of cosmic speed limit built into the structure of space and time, it makes sense to use it to **define** our unit of length (the meter) in terms of our unit of time (the second). This is what is done in SI, the International System of Units. It is no longer necessary to keep a "standard" 1 meter bar in a vault someplace. The second is now defined in terms of a particular oscillation of an atom in an atomic clock. The meter is then defined as the distance that light travels in 1/299,792,458th of a second. I should say that when I talk about the speed of light, I always mean the speed of light "in vacuum" — that is in empty space.[1] Things get more complicated in material like glass because the interactions of the light with the material can slow the light down.

Now 299,792,458 m/s is fast. It is a heck of a lot faster than we can actually move ourselves. But it is certainly not infinitely fast. With modern electronics, we can measure very short times, so it is not impossible to see the effect of the finite speed of light even over fairly short distances. The point I am trying to make here is that while motion at close to the speed of light is far beyond our everyday experience, it is not science fiction. In fact, we routinely measure the speed of light, and routinely see things (small things like electrons, but things nevertheless) moving at speeds very close to the speed of light.

But the surprising thing about light in a vacuum is that the speed of light that we measure doesn't depend on the velocity of the object that produced the light, and it doesn't depend on the velocity of the measuring apparatus. Now if you think that you understand this, you obviously have not been listening carefully enough, because this doesn't make any sense at all. Nevertheless, it is true. If, for example, I am running towards a light-bulb at speed v carrying a light-speed meter, a device to measure the speed of light, all of you sitting at rest see the light from the bulb approaching me at a speed $v + c$. But when I do the measurement, I get the same value for the speed of light that I would get if I were standing still. In fact, I get the same value that you would get measuring the same light beam in about the same place at about the same time, but standing still. The same thing happens if I am running away from the light source.



(1)

This is absolutely crazy. Surely if I am moving towards the light beam, I should register a larger speed on my light-speed meter. That is what common sense would say. However, that is not the way the world works. The way the world works is that the speed of light in vacuum is constant, period! It is not that something goes wrong with my light-speed meter. This bizarre fact is built into the way the world works.

The full power of this remarkable fact, the constancy of the speed of light, is unleashed when we combine it with another, much more reasonable fact about the way the world works — the principle of relativity. The principle of relativity says simply that all uniform motion is relative.

[1]The notion of "empty space" is itself rather problematic. Even classically, space is only completely empty at absolute zero. And when we include the effects of quantum mechanics, as we will see much later, empty space begins to look anything but empty. Nevertheless, there is a well-defined meaning to the notion of the speed in light in vacuum. Its role as a cosmic speed limit survives all this extra complication.

There is no absolute sense in which I can say I am moving. There is no preferred notion of standing still. In a moment, we will formalize this idea with the notion of an inertial frame of reference. Note that we can tell if our motion is not uniform. Acceleration is accompanied by forces that we can feel in our bones. But uniform motion is not detectable, so long as everything else we need is moving along with us. This, of course, **is** something that feel in our bones for the slow motions that we are used to. We all know this very well from travel in vehicles, cars, trains, planes, and whatnot. We are going to assume, with Einstein, that it remains true at relativistic speeds.

Inertial frames

The idea of an **inertial frame of reference** or just "inertial frame" for short, is one that already plays an important role in non-relativistic mechanics. It is an attempt to formalize the notion that motion is relative in an operational way. To do this, we must carefully describe what velocity means by describing precisely what we need to measure it.

On the surface, the speed of light does not seem to be a complicated concept. You measure it in the obvious way with clocks and meter sticks, by dividing the distance traveled by the time taken. But first, you have to synchronize your clocks! This is where the idea of an inertial frame comes in. An **inertial frame** is a real or imaginary collection of clocks that are fixed with respect to one another and synchronized, for example by requiring that some signal that originates midway between each pair of clocks arrives at the two clocks at the same time.[2] In addition, an **inertial frame** must not be accelerating, which is easy to check because you can just demand that Newton's laws hold for small velocities — free particles travel in straight lines, that sort of thing.

So we have two fundamental principles.

A. That the laws of physics are the same in all inertial frames, and

B. That one of the laws of physics is that the velocity of light is a constant — with the same value in all inertial frames,

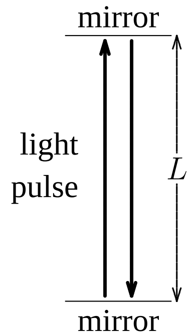
As you will see in more detail in the notes, these two principles are amazingly powerful. They will revolutionize our picture of space and time. Now let's see some of the consequences of putting these two ideas about the world together.

Time dilation

Let's start with one of the strangest and most trivial of the consequences of relativity — time dilation. The phenomena of time dilation can be stated precisely as follows. Observations done on a single clock moving with speed v with respect to a number of clocks fixed in an inertial frame show the ticking of the moving clock slowed down by a factor of $\sqrt{1-v^2/c^2}$. The standard way

[2] It doesn't matter for this purpose whether we are using a light signal or the Pony Express, as long as the signal travels at the same speed in both directions!

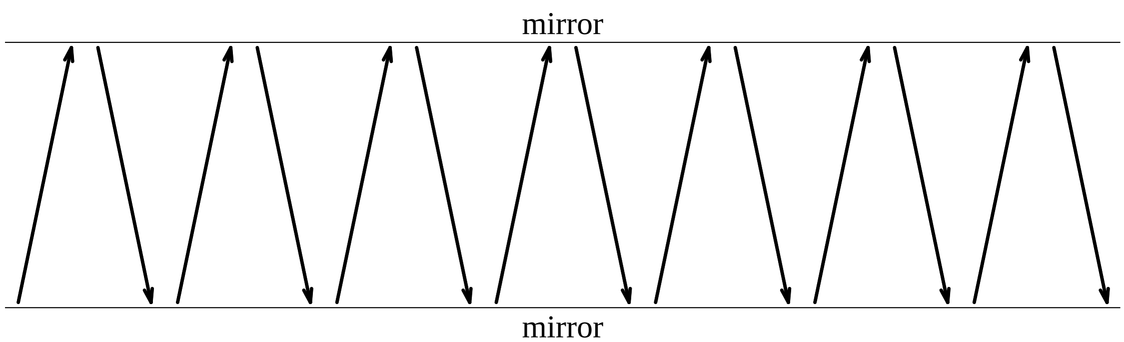
of deriving this result is to consider an idealized clock made out of two parallel mirrors and a pulse of light bouncing back and forth between them:



(2)

A light-clock at rest

If the distance between the mirrors is L, the time for each tick of the clock, defined as the time for the pulse to get from one mirror to the other, is L/c. Now suppose that the two mirrors of this light clock are mounted on parallel tracks a distance L apart and the two mirrors moved down the tracks with velocity v. Now the system looks to observers in the fixed frame as shown below



(3)

A light-clock in motion

Obviously, from the point of view of the many clocks in the inertial frame, the light pulse has to go farther when the single light clock is moving. Thus if light always travels at the same velocity, the ticks of the light clock take longer when it is moving. Call the factor by which the ticks are longer γ. Then we can compute γ as follows. Each vertical transit of the light from one mirror to the other in the moving frame takes time $\gamma L/c$ (just from the definition of γ - because L/c is the time for a tick in the rest frame). And because light travels at the same speed, c, the length of the path from one mirror to the next therefore must be γL. The light pulse moves vertically a distance L (because the tracks are a distance L apart) and horizontally a distance $v \cdot \gamma L/c$ (just the velocity of the light clock multiplied by the time). Now we look at the geometry of the motion. Then Pythagoras tells us that

$$(\gamma L)^2 = (\gamma L v/c)^2 + L^2 \tag{4}$$

which implies

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \tag{5}$$

So time is no longer sacred. And this can't just be a special property of light clocks, because if we used some other kind of clock to measure the time, and got a different result, then we would be able to distinguish between the moving frame and the frame in which the light clock is fixed. But this violates the principle that all frames are equivalent. Every kind of clock must tick out seconds at the same rate in all inertial frames.

Incidentally, this factor γ is going to reappear all the time, so it pays to actually either memorize it, or to be able to reproduce the light-clock argument in real time so you can get it whenever you need it.

It is quite easy with modern electronics and atomic clocks to see relativistic effects like time dilation. In fact, both special and general relativistic effects are very important in one very practical application — the Global Positioning System which is based on a system of atomic clocks aboard satelites. The relativistic corrections are small, because the satellites are traveling at "only" about 4000 m/s, but enormous accuracy is required to make GPS work and the relativistic effects must be properly included. See for example http://www.physicscentral.com/writers/2000/will.html.

Even more dramatic examples of time dilation occur all the time with elementary particles. That seems like a lot to ask of tiny particles that are supposed to be elementary and have no internal structure. But the fact is that quantum mechanics provides us with internal clocks for many elementary particles because they are unstable, and when they are sitting still and evolving in time, they have a constant probability per unit time of decaying into other lighter particles. We can actually see these internal clocks ticking (at least in an average sense) by watching the particles decay. The observed lifetime of unstable particles is a tangible measure of how fast these internal clocks are ticking. We see this all the time in particle experiments. But we are relying on another fact — all particles of a particular kind are exactly the same. We never actually measure the decay rate of the same decaying particle in two different frames. But we can quite easily measure the lifetime of a particle at rest, and then measure the lifetime of the same TYPE of particle in a moving frame. We find that the ratio of the lifetimes is γ. Since all particles of a particular type are identical, this is just as good.

Another important thing about time dilation is that although it is strange, it is probably the easiest of the relativity principles to remember and use. The thing to remember is

$$\boxed{\text{The single clock measures the shorter time.}} \tag{6}$$

If you keep this in mind, and just remember that $\gamma > 1$, you will always be able to reconstruct the right formula. Use time dilation whenever you can to solve problems!

The twin paradox

Now you may very well be thinking, at this point, that once you define what you are talking about carefully, with inertial frames, that there isn't anything particularly strange about motion at relativistic speeds, but that we have just confused the issue with a bizarre definition of measurement. Even our experiment on decays of elementary particles might be just a matter of a bad definition of what we mean by the ticking of their internal clocks. Perhaps, you think, that there is some other way of constructing our light-speed meters so that the speed of light is not constant and the bizarre features of relativity go away. Think again! Perhaps the simplest way of making clear that something totally bizarre is going on is to discuss the twin paradox. This is a classic thought experiment in which one twin takes a trip on a rocket moving at relativistic speeds, while the other twin remains at home. When the traveling twin returns, because his clocks have been ticking slowly, he is younger than his twin. His biological clock is no different from any other clock. Relativity has slowed down the aging process. If this is not strange, I don't know what is.

Again, this experiment has not been done conclusively with people. Astronauts in MIR or the space shuttle do age less rapidly than the rest of us, but the difference is sufficiently small at the speeds of mere orbital motion that the don't see a huge difference in biological clock (which aren't very accurate). The difference can be measured by atomic clocks. And a number of very accurate experiments have been done showing exactly this effect with the internal clocks of unstable elementary particles.

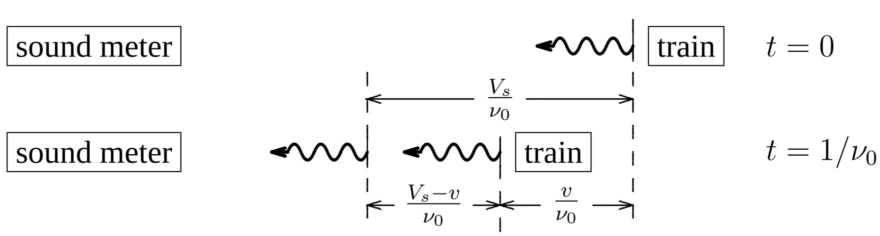
The twin paradox is so peculiar that I want to work out an example of how it looks to the two twins who are aging differently. To do this, it is useful to first understand the relativistic Doppler effect.

The Doppler effect

While the speed of a light beam does not change when we go from one inertial frame to another, its frequency does change. This is not surprising, since the same thing happens for sound or any other wave. It is called the Doppler effect and shows up, for example, in the change in sound of a train whistle when the train goes by. First, let me remind you how the Doppler effect goes for sound, or any other wave when we are moving at nonrelativistic speeds. Suppose that a train is moving towards me at speed v and its whistle emits sound waves which, when the train is at rest, have frequency ν_0, and wavelength λ_0. The speed of sound, V_s, is the product of the frequency and the wavelength:

$$V_s\,(\mathrm{m/s}) = \lambda_0\,(\mathrm{m/cycle}) \cdot \nu_0\,(\mathrm{cycles/s}) \tag{7}$$

The inverse of the frequency is the period of the sound wave, which is the time between successive crests of the wave. So now let us look at two successive crests of the wave as the train moves towards us — illustrated below:



(8)

Because the train is moving forward as it emits the wave, the crests are closer together than they would be if the train were standing still. The distance between crests for the train at rest is just the wavelength,

$$\lambda_0 = \frac{V_s}{\nu_0} \tag{9}$$

The distance between crests for the moving train is

$$\lambda_v = \frac{V_s - v}{\nu_0} \tag{10}$$

Thus the wavelength of the sound as recorded at the sound meter is reduced by the nonrelativistic Doppler factor

$$\frac{V_s - v}{V_s} \tag{11}$$

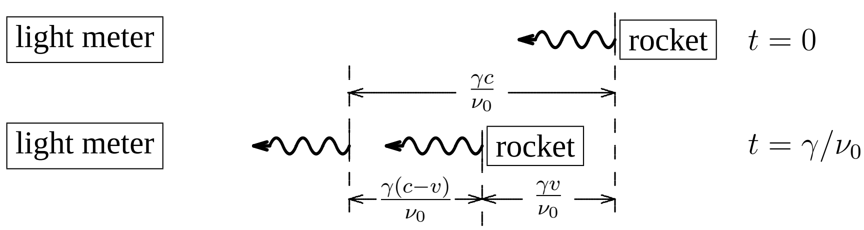
Because (7) must be satisfied, the frequency is increased by the inverse of (11), and the train whistle has a higher pitch when it is moving towards us.

If the train is moving away, the argument is exactly the same — we just have to replace $v \to -v$ in (11).

Now suppose we do a similar thing, but replace the train with a rocket moving at relativistic speed, and replace sound with light. The speed of light is the product of the frequency and the wavelength:

$$c\,(\mathrm{m/s}) = \lambda_0\,(\mathrm{m/cycle}) \cdot \nu_0\,(\mathrm{cycles/s}) \tag{12}$$

The inverse of the frequency is the period of the light wave, which is the time between successive crests of the wave. In the diagram, almost everything is the same, except that because of time dilation, the time between the emission of successive crests of the wave is longer than $1/\nu_0$ by the ubiquitous factor of γ, because the moving clock ticks more slowly. Thus the picture looks like



(13)

Because the rocket is moving forward as it emits the wave, the crests are closer together than they would be if the rocket were standing still. The distance between crests for the rocket at rest is just the wavelength,

$$\lambda_0 = \frac{c}{\nu_0} \tag{14}$$

The distance between crests for the moving rocket is

$$\lambda_v = \frac{\gamma(c - v)}{\nu_0} \tag{15}$$

Thus the wavelength of the light as recorded at the light meter is reduced by the relativistic Doppler factor

$$\frac{\gamma(c - v)}{c} = \frac{1}{\sqrt{1 - v^2/c^2}}(1 - v/c) = \sqrt{\frac{1 - v/c}{1 + v/c}} \tag{16}$$

Because (12) must be satisfied, the frequency is increased by the inverse of (16),

$$\sqrt{\frac{1+v/c}{1-v/c}} \tag{17}$$

and the light has higher frequency when the rocket is moving towards us. This is called blue-shift because raising frequency in the optical spectrum is a shift towards the blue.

Again, if the rocket is moving away, the argument is exactly the same — we just have to replace $v \to -v$ everywhere. This is called red-shift because lowering frequency in the optical spectrum is a shift towards the red.

There is one very important distinction to note about the relativistic Doppler effect versus the nonrelativistic version. In the relativistic version, it doesn't matter whether the rocket is approaching the observer at speed v or the observer is approaching the rocket at speed v. It can't, because of the principle of relativity. This is not true for the nonrelativistic Doppler effect because the air in which sound moves defines a special frame.

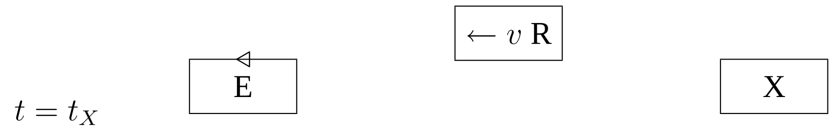
The twin paradox and inertial frames

We can now use the Doppler effect to understand time dilation and the twin paradox, by keeping track of every tick of the moving clock. So we have two twins, who are, as twins often are, the same age. But one twin (called the "rocket twin") takes a trip to planet X, at distance L from earth, staying in constant communication with the "earth twin" behind on earth. The communication is done using radio waves, or some other electromagnetic waves that have fixed frequency in the rocket twin's frame (the rocket frame). We will assume the typical form of the twin paradox that we discussed above, where the rocket twin goes out to planet X at speed v, quickly turns around, and returns at the same speed.

Here are a series of snapshots showing the important times during the trip. the rocket twin leaves the earth twin at $t = 0$ traveling at constant v – staying in constant radio communication

the rocket twin arrives at planet X at $t = L/v$, turns around (quickly) and sends a turn-around signal (shown as the open triangle $\triangleleft$ in the diagrams) to the earth twin, indicating that he has reached the planet.

The turn-around signal reaches the earth twin at $t = t_X = L/v + L/c$



the rocket twin and the earth twin are reunited at $t = 2L/v$

The radio transmitter is a clock – number of ticks (cycles) is $t\,\nu$ where ν is frequency of transmitter. Thus the number of ticks sent by the rocket twin is $T\,\nu$, where T is the total time he aged on the trip.

But the earth twin receives red shifted photons for a time $t_X = L/v + L/c$ and blue shifted for $2L/v - t_X = L/v - L/c$. Therefore, the number of cycles received by the earth twin is

$$\nu\sqrt{\frac{1-v/c}{1+v/c}}\left(\frac{L}{v}+\frac{L}{c}\right)+\nu\sqrt{\frac{1+v/c}{1-v/c}}\left(\frac{L}{v}-\frac{L}{c}\right) \tag{18}$$

$$= \nu\sqrt{\frac{1-v/c}{1+v/c}}\,L\left(\frac{1+v/c}{v}\right)+\nu\sqrt{\frac{1+v/c}{1-v/c}}\,L\left(\frac{1-v/c}{v}\right)$$

$$= \nu\,\frac{2L}{v}\sqrt{1-v^2/c^2} = \nu\,T \tag{19}$$

Thus

$$T = \frac{2L}{v}\sqrt{1-v^2/c^2} \tag{20}$$

Nonrelativistically, we would have expected

$$T = \frac{2L}{v} \tag{21}$$

Thus the earth twin “sees” rocket clocks ticking more slowly by a factor of

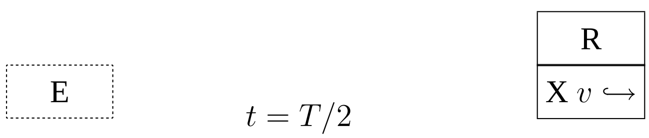
$$\sqrt{1-v^2/c^2} = 1/\gamma \tag{22}$$

This is time dilation. The moving clock ticks more slowly as seen by the clock at rest. This is also the twin paradox. Because the rocket twin has sent out fewer ticks, he has also aged less. He is younger than the earth twin when he returns.

There are a couple of other things to notice about (22). First note that the two terms in (22) are equal. This had to be the case, because the rocket twin sent the same number of ticks on the way to planet X as on the way back, so the earth twin received the same number of ticks in the red-shifted signal from the trip out as in the blue-shifted signal from the trip back.

But now, you say, why isn't the situation symmetrical? After all, from the rocket twin's point of view, the earth twin (along with the rest of the earth) has moved away at speed v and then come back at the same speed. Why is it that the rocket twin is younger at the end, rather than the earth twin? Let's look at the trip from the rocket twin's point of view, assuming that it is the earth twin who is sending out radio signals the whole time. Now things look a bit different. Here is the chronology. At $t = 0$, the rocket twin watches the earth twin and the earth recede at constant v. The rocket twin receives radio signals from the earth twin that are red-shifted until planet X appears. Planet X is shown as dashed in the figure because it is not in the same inertial frame as the rocket twin's ship, so the rocket twin has to be a whiz at relativity to calculate its position.

Planet X reaches the rocket twin at $t = T/2$, turns around (quickly) and starts to recede again. From this point on, the rocket twin receives blue-shifted signals until earth reappears and he is reunited with the earth twin.



the rocket twin and the earth twin are reunited at $t = T$

Now we can check that the two pictures are consistent. The number of cycles that the rocket twin receives is

$$\frac{T}{2}\nu\left(\sqrt{\frac{1-v/c}{1+v/c}}+\sqrt{\frac{1+v/c}{1-v/c}}\right)=\frac{T}{2}\nu\left[\sqrt{\left(\frac{1-v/c}{1+v/c}\right)\left(\frac{1-v/c}{1-v/c}\right)}+\sqrt{\left(\frac{1+v/c}{1-v/c}\right)\left(\frac{1+v/c}{1+v/c}\right)}\right] \tag{23}$$

$$\frac{T}{2}\nu\left(\frac{1-v/c}{\sqrt{1-v^2/c^2}}+\frac{1+v/c}{\sqrt{1-v^2/c^2}}\right)=T\nu\frac{1}{\sqrt{1-v^2/c^2}} \tag{24}$$

which using (20) is

$$\frac{2L}{v}\nu \tag{25}$$

which is the frequency times the time elapsed on earth, as we expected. I hope you can guess from this discussion what the asymmetry is. Of course I have not quite described this process the way the rocket twin experiences it. He doesn't just watch earth receding at $t = 0$. He blasts off and accelerates. Similarly, he doesn't just watch planet X turn around. He decelerates and accelerates again in the opposite direction. He feels these accelerations in his bones! From the rocket twin's point of view, the switch from red-shift to blue-shift occurs the moment he turns around. This makes sense. He knows that he has turned around. But from the earth twin's point of view, nothing special happens when the rocket twin reaches the planet. He has to wait until the turn-around signal arrives to see the shift from red-shift to blue-shift.

Here is another way of saying what the difference is. The earth twin, remaining on earth, is at rest in a single inertial frame the whole time. The rocket twin is not. He is in one frame on the way out, and in a different frame on the way back. It is the fact that the rocket twin must switch from one inertial frame to the other that makes his experience different.

Relativity may be strange, but it is consistent.

lecture 11

Topics:

- Where are we now?
- Space-time events
- The space-time interval
- The invariant interval
- Lorentz transformations
- Invariance as a way of thinking
- Varieties of space-time intervals
- Interlude on relativistic units
- The tip of tomorrow

Where are we?

We are lost in an odd 4-dimensional space! That's where! I hope the last lecture and the readings in Morin have convinced you that relativity is strange. We have discussed time dilation and the twin paradox in the last lecture. You should have read about the relativity of simultaneity and Lorentz contraction in David's book. By next week, I will be able to explain why I am so sure that relativity is true. For now, just try to accept that as strange as this is, there is absolutely convincing evidence that the world really works this way. Since we are stuck with it, we should figure out how to make sense of it, and how to avoid getting confused by the strangeness of it all. To do this, you must train yourself to consider space and time together, because what is time in one inertial frame is a combination of time and space in another. There are a number of closely related ways of dealing with this. I will talk today about a few things that you have already read about in David Morin's book: space-time events, the space and time intervals and the invariant combination of the two, and Lorentz transformations, because I want to try to put them in some perspective. At the end of this lecture, I will try to put a lot of this together and discuss in more detail the question of why is it that you are not used to thinking about space-time as a unified, 4-dimensional thing. Of course it is because you are slow - but there is more to say that I think is rather interesting.

Space-time events

The concept of a space-time event is very simple. Even in non-relativistic physics, to describe an event completely you must specify not only where it is but also when it takes place. **What's the big deal?** The importance of the space-time event in relativistic physics is really more of a negative one. There are many things that don't make **absolute** sense in the relativistic world because they look different in different inertial frames. Time is **relative**. Distance is **relative**. Simultaneity is **relative**. The space-time event gives us something to hang on to. The coincidence of space-time events is **NOT RELATIVE**. If space-time events coincide, that is if they happen at the same place and at the same time, they coincide in all inertial frames. Space-time events are thus the sensible things to talk about. Though the coordinates of a given space-time event will change depending on what inertial frame you're in, the event itself does not change.

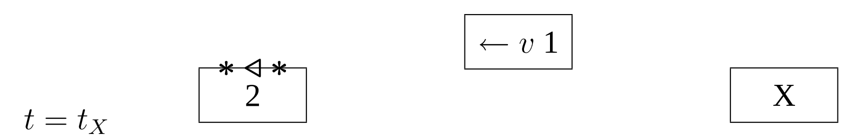
The right analogy to keep in mind is that different inertial frames in space and time are like different coordinate systems in three-dimensional space. Time and distance change in going from one inertial frame to another just as the x component of a point changes in going from one coordinate system to another. But space-time events in space and time are like points in three-dimensional space. If points coincide in one coordinate system, they coincide in all coordinate systems, just as space-time events are coincident in all inertial frames if the coincide in any one.

One of the things that we will try to train you to do is to identify the important space-time events in a process, and to be very suspicious of any statement or question that cannot be described in terms of space-time events. Let me illustrate the difference in the twin paradox discussion we gave in the last lecture, where twin 1 left twin 2 on earth and traveled to planet X and back. Here the crucial space-time events were:

Twin 1 leaves twin 2 — $\boxed{x = 0, t = 0}$ in Earth's frame, with Earth at the origin;

Twin 1 arrives at planet X — $\boxed{x = L, t = L/v}$ in Earth's frame;

The turn-around signal reaches twin 2 — $\boxed{x = 0, t = L/v + L}$ in Earth's frame;



Twin 1 and twin 2 are reunited — $\boxed{x = 0, t = 2L/v}$ in Earth's frame.

These events have an invariant meaning even though their coordinates, both their space and their time coordinates, will change if they are viewed in a different coordinate system.

It is perfectly meaningful to put together events, **so long as they occur at the same place and at the same time**. For example, again from the discussion of last lecture, "Twin 1 arrives at planet

X at $t = L/v$, turns around (quickly) and sends a turn-around signal to twin 2, indicating that he has reached the planet." This describes two separate events — reaching the planet and sending the signal — but because they happen at the same place and at the same time, their space and time coordinates will be the same in all reference frames. Thus the statement makes sense in all reference frames

If we don't stick to space-time events, there are lots of questions that we can ask that don't mean anything at all unless we specify the reference frame. For example[1]

Meaningless questions
How far does twin 1 travel?
How long does it take twin 1 to get to Planet X?
Where is twin 1 when the turnaround signal reaches Earth?
How old is twin 2 when twin 1 reaches planet X?

The last two are a classic kind of pitfall, because they attempt to compare two events at different places. Think about them carefully.

The space-time interval

Let me illustrate some of these issues by talking about a particularly useful sort of space-time event — the tick of a clock. If I have a clock sitting at the origin of the reference frame of the lecture hall, ticking every T seconds, the ticks of the clock define a series of space-time events. In the frame of reference of the lecture hall, all of these events have space coordinates 0, because they are all at the origin, and their time coordinates are just 0, T, $2T$, etc. Thus the time and space coordinates of the series of ticks would look like

event	E_0	E_1	E_2	$\cdots$
$(t, \vec{r})$	$(0, 0)$	$(T, 0)$	$(2T, 0)$	$\cdots$

(1)

In general, if we hadn't put the first tick at the origin, the $\vec{r}$ and t coordinates of the later ticks would all just have the initial $\vec{r}$ and t positions added on to them, so the sequence would look like

event	E_0	E_1	E_2	$\cdots$
$(t, \vec{r})$	$(t_0, \vec{r}_0)$	$(t_0 + T, \vec{r}_0)$	$(t_0 + 2T, \vec{r}_0)$	$\cdots$

(2)

We know that there is nothing very interesting about the $\vec{r}_0$ and the t_0, which can be changed by moving to a different coordinate system by translations in space and time. Thus it is useful to

[1]The word "meaningless" in the table is a little strong. It might be fairer to say "questions whose answers depend on the reference frame," but we are trying to get your attention!

get rid of this dependence. We can do this in a way that is completely independent of our choice of origin of coordinates by considering the **space-time intervals**, which are just the differences between the coordinates of space-time events

$$\Delta E_{01} = E_1 - E_0 = (\Delta t_{01}, \Delta\vec{r}_{01}) = (t_1 - t_0, \vec{r}_1 - \vec{r}_0) \tag{3}$$

intervals	ΔE_{01}	ΔE_{02}	$\cdots$
$(\Delta t, \Delta\vec{r})$	$(T, 0)$	$(2T, 0)$	$\cdots$

(4)

Events are the important underlying things, but have all this extraneous dependence on the origin of the coordinate system in space and time, so it is almost always useful to think about space-time intervals when you want to calculate something.

So far there is nothing relativistic about this, but here it comes. We also know from our discussion of time dilation what these same events look like in a reference frame in which the clock is moving. Suppose that the clock is moving with speed v in the $+x$ direction. Now because of time dilation, the ticks are spread out by a factor of γ, so that (if we set our clocks so that the first tick is at zero) the time coordinates are 0, γT, $2\gamma T$, etc. Let us also assume that the space coordinates are 0 for the first tick (at $t = 0$). Then because we know the velocity of the clock, we can easily work out the other $\vec{r}$ coordinates –

event	E_0	E_1	E_2	$\cdots$
$(t', \vec{r}\,')$	$(0, 0)$	$(\gamma T, \hat{x}v\gamma T)$	$(2\gamma T, 2\hat{x}v\gamma T)$	$\cdots$

(5)

Notice where the asymmetry of time dilation comes from here. We get these factors of γ for the clock readings in our frame because we are using a **series** of our clocks at different places in our frame to observe the ticks of the **single** moving clock. This is always the way time dilation works. **The single clock measures the shorter time!** We in our frame see the

Again, if we hadn't put the first tick at the origin in space and time, the x and t coordinates of the later ticks would all just have the initial $\vec{r}$ and t positions added on to them, so the sequence would look like

event	E_0	E_1	E_2	$\cdots$
$(t', \vec{r}\,')$	$(t'_0, \vec{r}\,'_0)$	$(t'_0 + \gamma T, \vec{r}\,'_0 + \hat{x}v\gamma T)$	$(t'_0 + 2\gamma T, \vec{r}\,'_0 + 2\hat{x}v\gamma T)$	$\cdots$

(6)

And again, we can get rid of the dependence on the origins in space and time by looking at the intervals:

intervals	ΔE_{01}	ΔE_{02}	$\cdots$
$(\Delta t', \Delta\vec{r}\,')$	$(\gamma T, \hat{x}v\gamma T)$	$(2\gamma T, 2\hat{x}v\gamma T)$	$\cdots$

(7)

Finally, there is nothing special about the x direction. For a different moving clock (or a different choice of coordinate system), the motion might be rotated into a different direction, $\hat{x} \to \hat{v}$

intervals	ΔE_{01}	ΔE_{02}	$\cdots$
$(\Delta t'', \Delta\vec{r}\,'')$	$(\gamma T, \hat{v} v\gamma T)$	$(2\gamma T, 2\hat{v} v\gamma T)$	$\cdots$

(8)

While you are used to thinking of the difference between (1) and (2) or between (7) and (8) as a change of coordinate system, you are probably not used to thinking about the difference between (1) and (6) as a change of coordinate system. But what you see here is that these space-time coordinates really can describe the same set of events, the ticks of a clock, in different reference frames. Going from one reference frame to another IS like going from one coordinate system to another — it is just that the time "coordinate" as well as the space coordinates are involved. The events don't change, but both the time and space coordinates change when we go from one inertial frame to another. What this means is that we have gone from a 3-dimensional space to a 4-dimensional space-time.

The invariant interval

We can extend the analogy between going from one coordinate system to another and going from one reference frame to another in the following way. When you make a change of coordinate system in space, the distance between points doesn't change. You can easily compute the square of the distance between two points labeled by vectors $\vec{r}_1$ and $\vec{r}_2$ in terms of their coordinates,

$$\ell^2 \equiv \left(\vec{r}_1 - \vec{r}_2\right) \cdot \left(\vec{r}_1 - \vec{r}_2\right) = \sum_{j=x,y,z} \left(r_1^j - r_2^j\right)^2 \tag{9}$$

Two things have happened here. First we have subtracted the coordinates of the two points. This eliminates dependence on the origin of the coordinate system, which just cancels when we subtract. Another way of putting this is that we are not interested in the length of the vectors themselves, because this depends on the arbitrary position of the origin. But if we make the vector from $\vec{r}_2$ to $\vec{r}_1$ by forming the combination $\vec{r}_1 - \vec{r}_2$, it doesn't depend on the origin. Then we sum the squares of the coordinates to get something that doesn't change when we make a rotation.

When you make a change of coordinate system in space-time, by going to a new reference frame, there is a similar thing that doesn't change. If I have two events, event 1 at time t_1 and position $\vec{r}_1$ and event 2 at time t_2 and position $\vec{r}_2$, the following combination doesn't change when we go from one reference frame to another:

$$s^2 \equiv c^2\left(t_1 - t_2\right)^2 - \left(\vec{r}_1 - \vec{r}_2\right) \cdot \left(\vec{r}_1 - \vec{r}_2\right) \tag{10}$$

doesn't change when we go from one reference frame to another. Again we have done two things. The obvious one is to subtract the space and time coordinates of the two events, so that s^2 doesn't depend on the origin of coordinates. The other, much less obvious one, is to combine the difference in time and the difference in space in a very particular way. You can see that this is the right thing to do by looking at two successive clock ticks in any of (4), (7) or (8), we have

$$s^2 = c^2T^2 - 0^2 = c^2T^2 \tag{11}$$

In (6), we have

$$s^2 = c^2\Big(\gamma T\Big)^2 - \Big(v\gamma T\Big)^2 = (c^2 - v^2)T^2/(1 - v^2/c^2) = c^2T^2 \tag{12}$$

This is independent of v, so it looks the same in any frame of reference.

The quantity s^2 in (10) looks a lot like the quantity ℓ^2 in (9). There are three differences. One is that time is involved. The second is the minus sign in front of the second term. The third is the factor of c^2 in the first term.

Let's deal with the easy one first. The factor of c^2 is nature's way of telling us that we are using a really stupid system of units. Because time and space get mixed up in relativity, and because the ratio v/c appears in many many of the relations of relativity, it makes sense to use units in which $c = 1$. One way of doing this is to use seconds as your unit of time and light-seconds, that is the distance light travels in a second, 299,792,458 meters. Of course, because we are so slow, this is an inconveniently long distance (which is why we don't adopt this convention as a standard part of our metric system) but you will get used to it. Anyway, 1 is easier to remember than 299,792,458. So from now on, we will follow nature's advice and set $c = 1$, which we will call "relativistic units." This will make our formulas look simpler. We also lose something by doing this, but it is not very important, and we will come back and discuss it at the end of the lecture today.

The other two differences between s^2 and ℓ^2 express the crucial strangeness of relativity. Time is not quite the same as space — that is obvious from the minus sign. But it is not completely different either, because it must be included to get something that looks the same in different reference frames.

Notice that because of the minus sign, s^2 can have either sign. It is positive in our example of the interval between two ticks of a clock, but for other kinds of intervals, such as the interval between two events that are in different places but at the same time in some frame, s^2 can be negative as well. The quantity s^2 goes by various names. I will call it the "invariant interval."

Lorentz transformations

You have read about a nice derivation of Lorentz transformation in David Morin's text. I will write them down and talk about them. They look much simpler with $c = 1$. So suppose that we have two events

$$\begin{aligned} &\text{event 1 at } t_1,\, \vec{r}_1 \text{ and} \\ &\text{event 2 at } t_2,\, \vec{r}_2. \end{aligned} \tag{13}$$

The interesting thing is the interval in time and space between these two events:

$$\Delta t = t_2 - t_1\,, \qquad \Delta\vec{r} = \vec{r}_2 - \vec{r}_1\,. \tag{14}$$

Now if we look at these same two events from a reference frame moving in the $+x$ direction with speed v, the coordinates of the events will change. We will get a new description of the events in terms of new coordinates,

$$\begin{aligned}&\text{event 1 at } t'_1,\, \vec{r}_1{}' \text{ and}\\ &\text{event 2 at } t'_2,\, \vec{r}_2{}'.\end{aligned} \tag{15}$$

Again, the interesting thing is the interval in time and space between these two events:

$$\Delta t' = t'_2 - t'_1 \qquad\qquad \Delta\vec{r}\,' = \vec{r}_2{}' - \vec{r}_1{}' \tag{16}$$

Now the statement of the Lorentz transformation is that we find that we can choose a coordinate system in which

$$\begin{aligned}\Delta x' &= \gamma(\Delta x - v\Delta t)\,,\\ \Delta t' &= \gamma(\Delta t - v\Delta x)\,,\\ \Delta y' &= \Delta y\,,\quad \Delta z' = \Delta z\,.\end{aligned} \tag{17}$$

If we set $\Delta\vec{r} = 0$ and look at $\Delta x'$ and $\Delta t'$ as a function of Δt, we can recognize the moving clock. The condition $\Delta\vec{r} = 0$ means that we are looking at a series of events that are fixed in the original frame. A pair of events with $\Delta\vec{r} = 0$ and time difference Δt could be two ticks of a fixed clock a time Δt apart.

$$\begin{aligned}&\Delta\vec{r} = 0\\ &\Delta x' = \gamma(-v\Delta t)\,,\\ &\Delta t' = \gamma(\Delta t)\,,\\ &\Delta y' = 0\,,\quad \Delta z' = 0\,.\end{aligned} \tag{18}$$

The γ in

$$\Delta t' = \boxed{\gamma}(\Delta t)$$

expresses time dilation. The $-v$ in

$$\Delta x' = \gamma(\boxed{-v}\Delta t)$$

describes the clock moving with x velocity $-v$ in the new frame.

The other terms in the Lorentz transformation can be understood in a similar way.

Notice that the Lorentz transformations are linear transformations. This is important for many reasons, but one of the most obvious is that the time evolution of a clock in some reference frame can be described by a straight line in space and time. The transformation from one frame to another must take any such straight path into another straight path.

It very important to understand why Lorentz transformations exist. They are a kind of space-time analog of rotations, but they look different because of the minus sign in (10). I recommend that you learn about Lorentz transformations, that you learn how to use them, but that you actually use them in a problem only if forced to do so (that is if the problem says "Use Lorentz transformations to ..."). There are usually easier ways of understanding any given problem. Lorentz transformation will get you there in the end if you don't make any mistakes, but you have to be very careful and systematic to use them properly.

Invariance as a way of thinking

Throughout this course so far, I have been subtly (and sometimes not so subtly) trying to get you to use symmetry arguments and invariance as a way of thinking about problems. In most of what we do, this is just a convenience - it makes hard problems easier. But in relativity, it is absolutely essential. The point of the principle of relativity after all is that inertial frames give equivalent descriptions of the physics. It is pretty obvious that you should use this fact whenever you can to understand what is going on. That is why so many of the problems you will do about special relativity are much easier when you choose some particularly convenient inertial frame and/or compute some appropriate invariant quantity. I want you to understand special relativity as much more than a confusing collection of formulas. Even though you will never learn to feel special relativity in your bones, you can learn to appreciate it as an interconnected web of relations between different ways of looking at the same phenomena.

While Lorentz transformations are often not the best way to solve problems, they will nevertheless be very important to our understanding of relativity. Lorentz transformations are part of the basic symmetry of space and time, along with rotations in space and translations in time and space, and the invariant interval is invariant under all these transformations. The symmetry transformations and the invariant quantities that they leave unchanged are two sides of the same coin. We will very often use the fact that the invariant interval looks the same in all inertial frames. And we will find other invariant quantities as well. So while we will not always be using Lorentz transformations explicitly, they will be with us implicitly because we will often be using the quantities that they leave invariant, like the invariant interval.

Varieties of space-time intervals

We spent a lot of time talking about the space-time interval $(\Delta t, \Delta\vec{r})$ between two ticks of a moving clock. This is called a "time-like" interval, because there is a frame in which it has a time component but no space component — the frame in which the clock is at rest. In fact, any space and time interval with components

$$(\Delta t, \Delta\vec{r}) \tag{19}$$

that satisfies

$$\Delta t > \left|\Delta\vec{r}\right| \tag{20}$$

has the property that there is a frame in which it looks like $(\Delta\tau, 0)$ with $\Delta\tau > 0$. Why? One way to see this is to note that any interval that satisfies (20) could be the interval associated with a clock moving with velocity

$$\vec{v} = \frac{\Delta\vec{r}}{\Delta t} \tag{21}$$

simply because if the clock moves with $\vec{v}$ for a time Δt, it ends up translated by the vector $\Delta\vec{r}$, reproducing (19). Thus if we go to a frame moving with velocity $\vec{v}$, we will be moving along with the clock and in this new frame the intervals will have the form

$$(\Delta\tau, 0) \qquad \text{where} \qquad \Delta\tau = \sqrt{\left(\Delta t\right)^2 - \left|\Delta\vec{r}\right|^2} \tag{22}$$

This relation incorporates the statement of time dilation. $\Delta\tau$ is always less than or equal to Δt.

Two events that are separated by a time-like interval are closest together in time in the frame in which they occur at the same point in space. (23)

The frame in which the two events occur at the same point in space is the frame in which the two events could be two ticks of the same clock. Thus (23) is equivalent to the statement that a moving clock ticks slowly. Two events that are separated by a time-like interval are sometimes referred to as "time-like separated."

On the other hand, if

$$\Delta t < \left|\Delta\vec{r}\right| \tag{24}$$

then it cannot represent the interval between two points on the path of a clock. As we will see in more detail later, we cannot go to a frame moving with velocity $\vec{v} = \Delta\vec{r}/\Delta t$, because $|\vec{v}\,| > 1$, so the frame would have to moving at a speed greater than the speed of light, which is impossible for clocks and meter sticks and all the other things we need to have a frame of reference. Then what is this interval? Suppose that we look in a frame of reference moving with velocity

$$\vec{u} = \Delta\vec{r}\,\frac{\Delta t}{\left|\Delta\vec{r}\right|^2} \tag{25}$$

To see what happens, let us rotate our coordinate system until $\Delta\vec{r}$ is in the x direction

$$\Delta\vec{r} = \left|\Delta\vec{r}\right|\hat{x} \tag{26}$$

so that

$$\vec{u} = u\,\hat{x} \qquad \text{where} \qquad u = \frac{\Delta t}{\left|\Delta\vec{r}\right|} \tag{27}$$

Now in a frame moving with velocity $\vec{u}$, the y and z components of the interval remain zero, and we can compute what the time and x components become using the Lorentz transformation,

$$\begin{aligned}
\Delta t' &= \frac{1}{\sqrt{1-u^2}}\left(\Delta t - u\left|\Delta\vec{r}\right|\right) = 0 \\
\Delta\vec{r}\,' &= \frac{1}{\sqrt{1-u^2}}\left(\left|\Delta\vec{r}\right| - u\,\Delta t\right) \\
&= \frac{1}{\sqrt{1-u^2}}\left|\Delta\vec{r}\right|\left(1 - u\,\frac{\Delta t}{\left|\Delta\vec{r}\right|}\right) \\
&= \frac{1}{\sqrt{1-u^2}}\left|\Delta\vec{r}\right|\left(1-u^2\right) \\
&= \sqrt{1-u^2}\left|\Delta\vec{r}\right| = \sqrt{\left|\Delta\vec{r}\right|^2 - \left(\Delta t\right)^2}
\end{aligned} \tag{28}$$

This is called a "space-like" interval, because there is a frame of reference in which it has a space component but no time component. A very important example of a space-like interval is the interval between two ends of a measuring stick at fixed time. This is what we define to be a measurement of distance. Two objects are a distance one meter apart in a given reference frame at time t if we can line up a meter stick so that the objects are at opposite ends of the stick at time t. This distance is associated with a space-time interval with is the difference between two "events" which are the space-time coordinates of the two ends of the meter stick at time t. This is obviously a space-like interval, because the time component is zero. This corresponds to the primed frame in (28). The properties of space-like intervals are then responsible for the phenomenon of "Lorentz contraction". In any other frame, the two events occur at different times. Therefore, (28) implies that $\left|\Delta\vec{r}\right| > \left|\Delta\vec{r}'\right|$. The two events are closest together in the frame in which they occur at the same time. This is generally true for any space-like interval.

Two events that are separated by a space-like interval are closest together in space in the frame in which they occur at the same time. (29)

Thus for example, if we measure the length of a moving train, the length we measure is the distance between two events describing the positions of the front and back of the train at the same time in our frame. But in the train frame, these two events do not occur at the same time, and thus the distance between them is greater than what we measure (by a factor of $1/\sqrt{1-v^2}$). This is Lorentz contraction.

Two events that are separated by a space-like interval are sometimes referred to as "space-like separated."

Interlude on relativistic units

Let's talk about the addition of velocity formula you saw in Morin

$$\frac{v_1+v_2}{1-v_1v_2/c^2}\,. \tag{30}$$

In sensible units in which $c=1$, this looks like

$$\frac{v_1+v_2}{1-v_1v_2}\,. \tag{31}$$

One can ask, suppose, in sensible units, you get a result like (31). How would you know how to put the factors of c back into (31) to get (30)? This is a crucial step because it allows you to do calculations with $c=1$, to make things simple, but still get the full answer at the end of the day, so you can talk to people who insist on using dumb units.

If you have done a calculation in relativistic units and you want to put the factors of c back in to translate the result to conventional units, you have to know the conventional units of the objects you have calculated. Then you put in factors of c to get the units right.

In the case of (31), we know that we are calculating a velocity, and the numerator is already a velocity. So we can leave that alone and make the denominator dimensionless. The 1 term is OK, and we can make the uv term dimensionless by dividing by two factors of c — and that is where (30) comes from.

Another example: Suppose that you are calculating an energy in terms of a mass m and a velocity v and you get the result

$$6\pi\, m\, v^4 \tag{32}$$

The 6π is a dimensionless number — we can forget about that. But the rest would not have units of energy in conventional units (remember, for example, that Newtonian kinetic energy is $\frac{1}{2}mv^2$). We have two too many factors of v, so we must divide by c^2 — thus the result in normal units is

$$6\pi\, m\, v^4/c^2 \tag{33}$$

Here's another one. For a system of two particles, with masses m_1 and m_2 and speeds v_1 and v_2, suppose you calculate an object that is a mass, and the result is

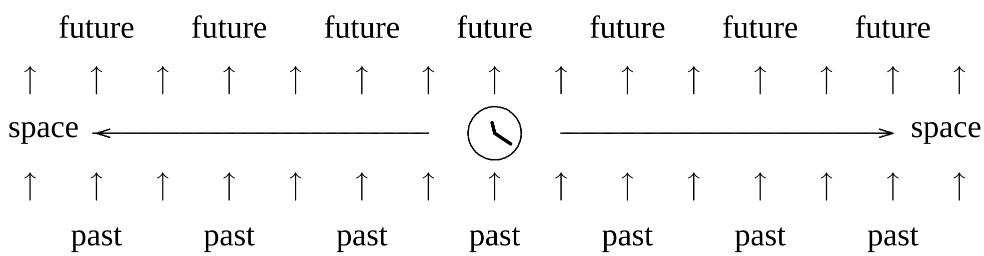
$$\frac{m_1/v_1 + m_2 v_2}{m_1/m_2 + v_1 v_2} \tag{34}$$

Let us start with the denominator. We don't know what the units should be of the denominator alone (because we can always multiply numerator and denominator by a power of c), but it is clear that the two terms in the denominator must have the **same** dimension. So for example, we can divide the second term by c^2. That makes the whole denominator dimensionless, so the numerator must then have the dimensions of mass. We can arrange this by multiplying the first term by c and dividing the second by c — thus the result in normal units is

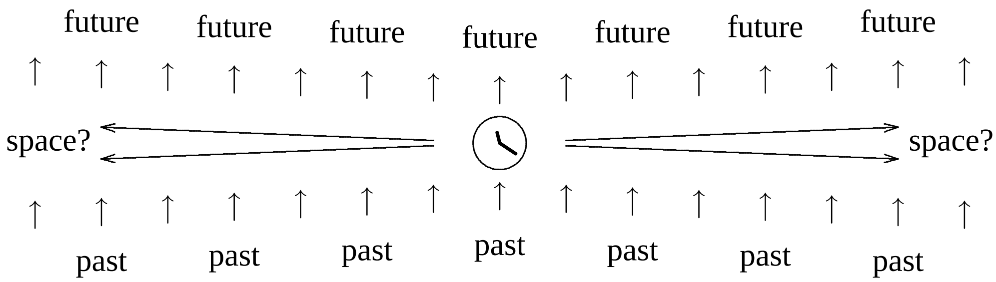
$$\frac{m_1 c/v_1 + m_2 v_2/c}{m_1/m_2 + v_1 v_2/c^2} \tag{35}$$

The tip of tomorrow

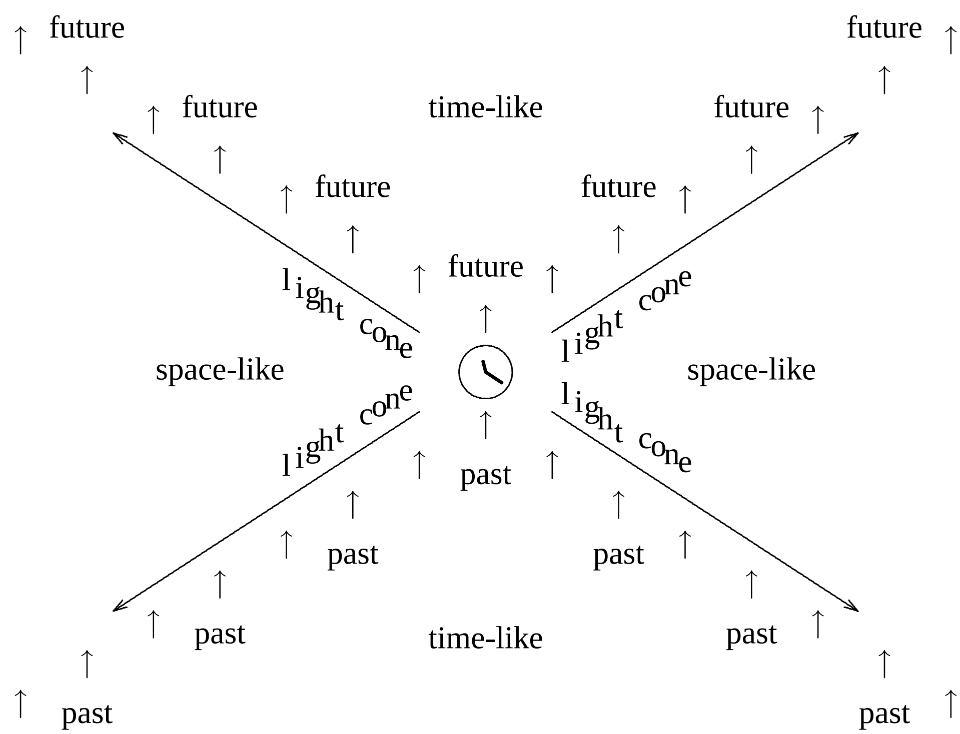
So then why is it that you are not used to thinking about space-time as a 4-dimensional object? I think that once again it is because you are slow, of course. But let's look at this in a little more detail in the light of space-time events and the varieties of space-time intervals. You are used to thinking of time and space very differently. Space is spread out around you to infinity in all directions. Time is not spread out at all. Time passes, and you live in the present on your way from the past to the future. Here is a cartoon of our subjective picture of space and time, plotted just in the $x - t$ plane.



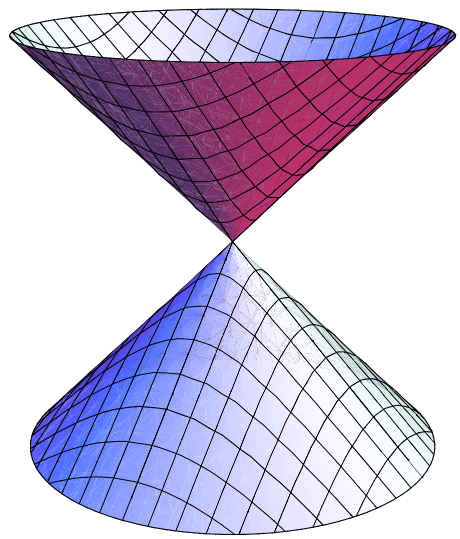
But much of the difference between space and time is an illusion. Time and space are different, to be sure, because of that crucial minus sign in the invariant interval. The passing of time is real. But the spreading out of space is a fiction. It is a drastic oversimplification that has been built into our brains and our sensory apparatus over thousands of generations because we are slow. What we don't think about or notice is that when we look out at the space spread out around us, we are actually looking back into the past. And we cannot access the future around us instantaneously. The real story looks more like this.



The slight slope of the long almost horizontal arrows is the non-zero value of $1/c$ - the effect of the finite speed of light. If we were faster, we would recognize intuitively that the "space" that is spread out on all sides of us between these arrows at this moment in time is just as inaccessible to us as the future. This is the region of space-time events from which we are separated by space-like intervals. We can't get to these events, nor can they get to us. This region is neither the past nor the future. There are three different regions of space-time - the past - the future - and the space-like separated that is neither past nor future. If we were fast, then the picture would look more like this.

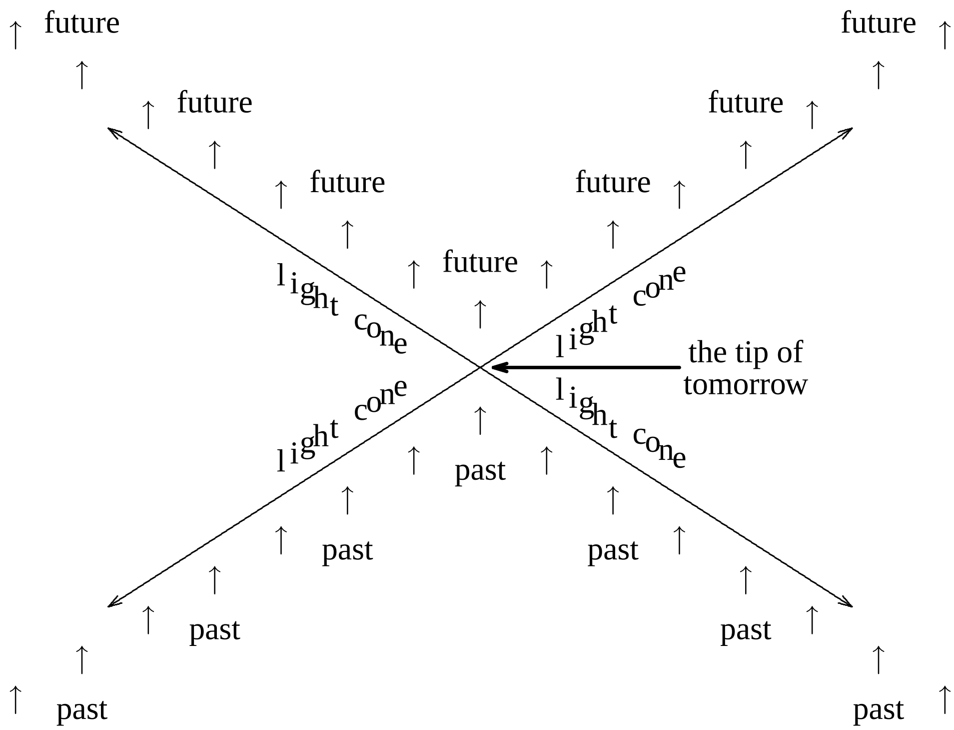


The boundary between the space-like separated region and the time-like separated regions of our past and future is a cone called our light cone. Our light cone consists of all the events in space-time that we could communicate with or receive communication from by means of a single light signal, moving at the speed of light. In the figure, it looks like two crossed lines, but we can rotate in space. If we rotate it just in the $x - y$ plane and keep $z = 0$ (for example), we get an ordinary three dimensional cone in the three dimensions, x, y and t. The light-cone of a given event, and with it the division of space-time into three regions - past, future, and space-like separated - looks the same in every reference frame because the speed of light does not change. The light cone is a funnel from our past to our future. In natural units with $c = 1$, it looks like this:



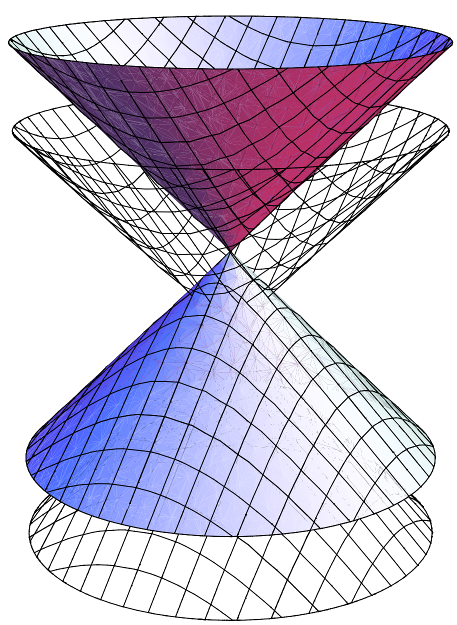
Of course, if we make an arbitrary space rotation, we get a four dimensional cone, which is harder to visualize, so I recommend just ignoring z for now.

The light-cone separates space-time into three regions, the time-like past which could have affected us, the time-like future that we can affect, and space-like unknown on all sides that can neither have affected us nor be affected by us. This picture doesn't matter much to us in our everyday life, because we are so slow. But it will be crucial when we try to understand the large-scale structure of the universe later in the course.

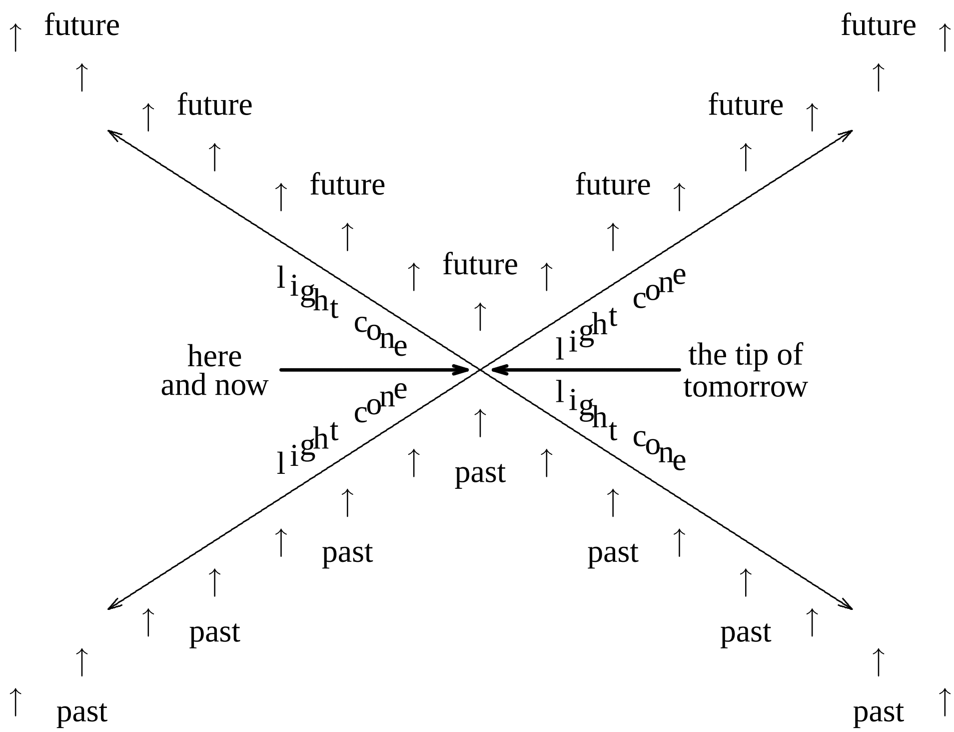


You can see in these pictures that our present is the singular tip where the past light cone meets the future light cone. This is what I call “the tip of tomorrow” because it is the single point in space-time that separates our past from our future.

As our present evolve into our future, our light-cone moves with us, further narrowing the future, and opening us to effects from events that were space-like separated.



For now, I hope that this picture helps you to understand why relativity seem so strange. And maybe this will help you to avoid the wrong thinking than can so easily get you derailed thinking about relativistic trains. Seen properly, through the lens of relativity, the “the present” is not in any sense an infinite three dimensional space. The present is both right NOW and RIGHT HERE!



lecture 12

Topics:
World lines and proper time
Massive particles
Energy and momentum
Example of inelastic process
Lorentz transformation of energy-momentum
Energy, momentum, velocity and mass
Massless particles
4-vectors and the invariant product

World lines and proper time

There is a nice relativistic analog of the concept of a **trajectory** in Newtonian physics - based on the concept of a **world line**. The idea is that as any massive particle evolves with time, whether it is sitting still or moving, free or accelerating, it can always be described by some curve in 4-dimensional space-time that is just the collection of the space-time events $\left(t, \vec{r}(t)\right)$ that describe where the particle is at every time. The reason that it is useful to think about this in this funny 4-dimensional space (even though it is kind of hard to visuallize) is that the world line itself has an invariant meaning because it is a collection of space time events that have invariant meaning even though their coordinates will change depending on the inertial frame. The situation is similar to that of a curve in three dimensional space. The curve consists of points that have an invariant meaning, but their coordinates change depending on the coordinate system. Also, just as we can label the points on a curve in three dimensional space by the distance along the curve from some arbitrary point, so we can measure the distance along world line of a particle by measuring the time ticked on the particles internal clock. This is called the proper time, τ. The change in proper time $d\tau$ along any short segment of the world line is just

$$d\tau = dt/\gamma = dt\sqrt{1-v^2} = \sqrt{dt^2 - d\vec{r}^2} \tag{1}$$

where t and v are measured in whatever coordinate system and inertial frame you like. You can see that this is independent of the inertial frame and the coordinate system by looking at the last term, which is simply the invariant interval for the short line segment. It is crucial here that every short line segment along the world line of a massive particle is time-like. This is true because massive particles can never travel as fast as light. We will return to this again and again and understand it various ways.

A good way of describing the world line is to give both t and $\vec{r}$ as functions of τ:

$$t(\tau) \qquad \vec{r}(\tau) \tag{2}$$

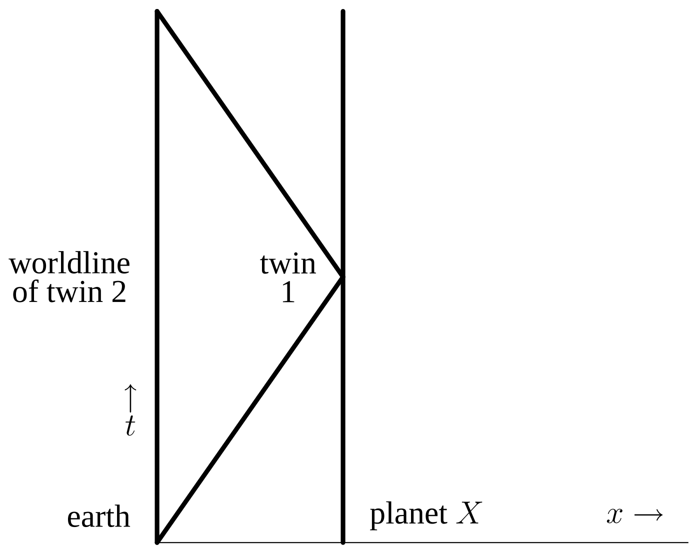
For example, for a particle at rest

$$t(\tau) = \tau \qquad \vec{r}(\tau) = 0 \tag{3}$$

and for a particle with velocity $\vec{v}$

$$t(\tau) = \gamma\tau \qquad \vec{r}(\tau) = \gamma\vec{v}\tau \tag{4}$$

For travel in one dimension, we can plot world lines on a plane. For example, consider the twin paradox. In their own frame of reference, the world lines of earth (and twin 2 on earth) and planet X are vertical lines. The world line of twin 1 in the rocket has a kink where the rocket turns around - in the curious geometry of spacetime - the kinked line takes the shorter time.



Massive particles

I now want to change gears from talking about space and time to talking about energy and momentum. I am going to write down the Lagrangian for a free relativistic massive particle. Because this is a free particle, the Euler-Lagrange equation is not all that interesting. We know without thinking about it that it is going to tell us that the particle moves at constant velocity. Nevertheless, finding the right form for the Lagrangian can be instructive. It will allow to ask questions about what would happen if we put in forces, for example. Also, it allows us to identify the quantities that we expect to be conserved because of Noether's theorem and translation invariance in time and space. In fact, once we write down these quantities — better known as the energy and the momentum — we will largely forget about the Lagrangian formulation. So don't panic if this Lagrangian looks so strange that you can't deal with it. What will really matter is that you learn to deal with the relativistic energy and momentum that come out of it from Noether's theorem.

What principles should we use to write down the Lagrangian for the free massive particle moving at relativistic speeds? Surely one important principle is that the Action should look the same in all reference frames. This is reasonable, because if it were not true, Hamilton's principle would not necessarily give us equations of motion that give the same trajectories in all reference frames. In addition, if we call the position of the particle $\vec{r}$, we expect that the Lagrangian is independent of $\vec{r}$ and of t, and depends only on $\dot{\vec{r}}$, because it should be invariant under translations in space and time. In a moment, I will show that the following Lagrangian gives rise to an action that is independent of the reference frame:

$$L(r,\dot{r}) = -m\sqrt{1-\dot{\vec{r}}^{\,2}}\,. \tag{5}$$

We are using units with $c = 1$ as usual. Here are a few things to note. This looks a bit funny because of the square-root. But as we will see, that is what relativity requires us to write down, so we will just have to live with it. The constant m, which we will see in a moment is the mass of the particle, must be there so that the Lagrangian has units of energy.

Now let us show what happens to the action

$$S = -m\int dt\,\sqrt{1-\dot{\vec{r}}^{\,2}} \tag{6}$$

under a Lorentz transformation. This is a little complicated because both the integrand and the dt change under a Lorentz transformation. But we can make what is going on more obvious by writing (6) as follows:

$$S = -m\int dt\,\sqrt{1-(d\vec{r}/dt)^2} = -m\int \sqrt{dt^2-(d\vec{r}\,)^2} \tag{7}$$

The second form is a very funny looking integral, but it makes sense because it is equivalent to the previous form. The important point is that the last form is evidently unchanged by Lorentz transformations. The infinitesimal interval

$$(dt, \vec{dr}) \tag{8}$$

is just a space-time interval. It transforms under Lorentz transformations just like $(\Delta t, \Delta\vec{r})$. And therefore, the combination

$$dt^2 - (d\vec{r}\,)^2 \tag{9}$$

is just an infinitesimal version of the invariant interval, and it has the same value in all inertial frames. To get something proportional to dt (so that we can put it under an integral sign and get a finite result), we must take the square-root of (9). That is why (6) looks the way it does, nutty as that is.

Energy and momentum

Now that we have a Lagrangian, we can construct the conserved energy and momentum that we expect for a relativistic particle. The energy is

$$E = \dot{\vec{r}}\cdot\frac{\partial L}{\partial\dot{\vec{r}}} - L = m\frac{\dot{\vec{r}}^{\,2}}{\sqrt{1-\dot{\vec{r}}^{\,2}}} + m\sqrt{1-\dot{\vec{r}}^{\,2}} = \frac{m}{\sqrt{1-\dot{\vec{r}}^{\,2}}} = \frac{m}{\sqrt{1-v^2}} = m\gamma\,. \tag{10}$$

The momentum is

$$\vec{p} = \frac{\partial L}{\partial\dot{\vec{r}}} = m\frac{\dot{\vec{r}}}{\sqrt{1-\dot{\vec{r}}^{\,2}}} = \frac{m\vec{v}}{\sqrt{1-v^2}} = m\vec{v}\gamma\,. \tag{11}$$

These things look really funny, but these are the energy and momentum that we expect to be conserved because of Noether's theorem. Of course, for a single free particle, this is pretty trivial. Nothing changes, so everything is conserved. But in fact, these objects (10) and (11) are conserved in exactly the way non-relativistic energy and momentum are conserved at low speeds. If we have a system with several particles and we add up all the energies or all the momenta, we get something that doesn't change as the system evolves in time. The energy and momentum may flow around from one particle to another, but the total does not change.

Here is an example. Suppose we have a particle with mass m traveling at $v = \frac{4}{5}$. Then

$$\gamma = \frac{1}{\sqrt{1-(4/5)^2}} = \frac{5}{\sqrt{5^2-4^2}} = \frac{5}{\sqrt{25-16}} = \frac{5}{\sqrt{9}} = \frac{5}{3} \tag{12}$$

so

$$E = \gamma m = \frac{5}{3}m \qquad\qquad p = v\gamma m = \frac{4}{3}m \tag{13}$$

You will observe that these do not look like the expressions for energy and momentum of a Newtonian particle. Nevertheless, these are the objects that are really conserved. It is instructive to look at them with the factors of c put back in:

$$E = \frac{mc^2}{\sqrt{1-v^2/c^2}} \tag{14}$$

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1-v^2/c^2}} \tag{15}$$

In this form, it is also useful to Taylor expand in powers of v/c in order to see what they look like at small velocity:

$$E = \frac{mc^2}{\sqrt{1-v^2/c^2}} = mc^2 + \frac{1}{2}mv^2 + \cdots \tag{16}$$

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1-v^2/c^2}} = m\vec{v} + \cdots \tag{17}$$

As we expected, the Newtonian expressions for the kinetic energy and the momentum appear. But in the expression for energy, (16), there is also a constant term, the famous mc^2. The mc^2 in (16) is a constant that is irrelevant to the Euler-Lagrange equations, so at small velocities, the velocity dependent part of (16) reduces to the non-relativistic Lagrangian, as we expect. What is actually important about the extra term is that mass plays a very different role in relativistic processes than it does at lower speeds. In Newtonian mechanics, mass, kinetic energy, and momentum are separately conserved. But in relativistic physics, it is only the momentum $\vec{p}$ and the total energy E that are actually conserved in arbitrary collisions of elementary particles at any speed — even in processes in which particles are created or annihilated. Mass is not conserved. The mass of any one kind of particle is always the same, so mass is conserved in collisions that don't change the type of particle involved. For example if you accelerate an electron to speed v and it collides with an electron at rest, some of the time you will get a collision in which you end up with two electrons going off in different directions and no other particles. We might represent this process by the schematic "equation"

$$e^- + e^- \to e^- + e^- \tag{18}$$

In this case the sum of the masses of the particles before the collision is the same as the sum of the masses of the particles after the collision. This is the analog in the relativistic world of an "elastic" collision in nonrelativistic physics, and I will sometimes use the same word to describe it. But other things can happen that don't conserve mass. For example, in the process of the collision, you may produce an extra electron (so there are three electrons in the final state) and a positron (an anti-electron - with the same mass but opposite charge).

$$e^- + e^- \to e^- + e^- + e^- + e^+ \tag{19}$$

Here the sum of the masses before the collision is $2m_e$, and the sum after the collision is $4m_e$. Mass is not conserved. In fact, there is no reason to compute the sum of the masses at all. It is just not an interesting quantity. In fact, the words "elastic" and "inelastic" mean something a little different when applied to relativistic collisions than the do they do in the nonrelativistic case. For a nonrelativistic elastic collision, the total kinetic energy, $\sum \frac{1}{2}mv^2$, is conserved. For a relativistic elastic collision, the total relativistic energy, $\sum m\gamma$, is conserved. In both cases, the particles in the initial state are the same as the particles in the initial state. For a nonrelativistic inelastic collision, kinetic energy is not conserved. Some of the kinetic energy is converted to heat and effectively lost. At the microscopic level, this is really a very complicated process in which kinetic energy of the inital objects is converted to kinetic energy associated with random motion of their parts. In relativistic physics, we use the term "inelastic" to mean something quite different — that new particles are created or destroyed. In this case, while energy is conserved, mass is not. This is summarized in the following table.

	small v	large v
elastic	$KE = \sum \frac{1}{2}mv^2$ is conserved — mass is conserved —	$E = \sum m\gamma$ is conserved particles are conserved therefore — mass is conserved —
inelastic	$KE = \sum \frac{1}{2}mv^2$ not conserved — mass is conserved —	$E = \sum m\gamma$ is conserved but new particles are produced or destroyed — mass is **not** conserved —

So once again, this shows that conservation of mass is just not something fundamental. Sometimes mass is conserved. Sometimes it isn't.

Example of inelastic process

Here is an interesting and important particle decay process

$$J/\psi \to e^- + e^+ \tag{20}$$

where e^- is an electron, e^+ is a positron, the antiparticle of the electron, and J/ψ is a particle that has two names for historical reasons. It was discovered allegedly independently at Brookhaven and at SLAC. Anyway, it has a mass of $m_{J/\psi} \approx 3097$ MeV. You can make it in particle collisions and it very quickly decays in one of various ways. Let's summarize what energy and momentum conservation implies for J/ψ at rest decaying into an electron and a positron.

particle	J/ψ	e^+	e^-
$\vec{v}$	0	$\vec{v}_+$	$\vec{v}_-$
E	$m_{J/\psi}$	$m_e\gamma_+ = m_{J/\psi}/2$	$m_e\gamma_- = m_{J/\psi}/2$
$\vec{p}$	0	$m_e\gamma_+\vec{v}_+$	$m_e\gamma_-\vec{v}_-$

(21)

$$\begin{aligned} &\text{energy conservation} \quad \Rightarrow \quad m_{J/\psi} = m_e\gamma_+ + m_e\gamma_- \\ &\text{momentum conservation} \quad \Rightarrow \quad 0 = m_e\gamma_+\vec{v}_+ + m_e\gamma_-\vec{v}_- \end{aligned} \tag{22}$$

$$\gamma_-\vec{v}_- = -\gamma_+\vec{v}_+ \quad \Rightarrow \quad \vec{v}_- = -\vec{v}_+ \tag{23}$$

Lorentz transformation of energy-momentum

One of the important things about energy and momentum is that they behave under Lorentz transformations very much like a space and time interval. Remember that this has to do with what happens to their values when we go from one reference frame to another. The easiest way to understand what happens to energy-momentum is to imagine that the particle has a clock on it and consider the space-time interval between two ticks of the particle's clock, as we did in the last lecture. Consider, then, the space-time interval between two ticks of a particle's clock. A space-time interval between two events has a time component, Δt, that is the time that elapses between the two events, and a space component, $\Delta\vec{r}$, that is the vector from the position of one event to the position of the other. In relativistic units, of course, these two components have the same dimension. If the particle is sitting still, this interval has a time component, call it $\Delta\tau$, but its space component vanishes.

Now suppose that the particle is moving with velocity $\vec{v}$. Then, because of time dilation, the time interval between the same two ticks of the particle's clock is

$$\Delta t = \Delta\tau \, \frac{1}{\sqrt{1 - v^2}} \tag{24}$$

But then, because

$$\vec{v} = \frac{\Delta \vec{r}}{\Delta t} \tag{25}$$

we find

$$\Delta \vec{r} = \Delta \tau \, \frac{\vec{v}}{\sqrt{1 - v^2}} \tag{26}$$

But then we can write

$$(E, \vec{p}) = \frac{m}{\Delta \tau}(\Delta t, \Delta \vec{r}) \tag{27}$$

But both m and τ are invariants — constants — they are properties of the particles involved, not of the frame. Thus we conclude that $(E, \vec{p})$ must transform just like $(\Delta t, \Delta \vec{r})$ because the two are just proportional to one another.

4-vectors and the invariant product

I hope that by this time you are getting used to the idea of changes from one inertial frame to another, and the accompanying Lorentz transformation, as a kind of 4-dimensional analog of what we do in three dimensional space when we go from one coordinate system to another. After developing the analogy between 3-dimensional space and 4-dimensional spacetime, how can we not go on to develop the analogy between 3-dimensional vectors and 4-dimensional vectors. Indeed, our treatment of the invariant interval was the first step in developing this analogy. We saw there that the invariant interval (this time with $c = 1$),

$$s^2 \equiv \left(t_1 - t_2\right)^2 - \left(\vec{r}_1 - \vec{r}_2\right) \cdot \left(\vec{r}_1 - \vec{r}_2\right) \tag{28}$$

is a kind of length. Like the length of a 3-dimensional vector, it has the same value in all coordinate systems, but here of course, the idea of coordinate system is enlarged to include different inertial frames, moving with different velocities.

The obvious way to go further is the find analogs in spacetime for the 3-vector, $\Delta \vec{r}$, and for the dot product. The analog of the 3-vector is pretty obvious. It is a 4-vector, with 4 components, the first of which (which we will call the t component, just to remind us that it is special) is the time. So a 4-vector is a quartet of numbers,

$$A = (A_t, A_x, A_y, A_z) \tag{29}$$

where A_t is refered to as the time component, and A_x, A_y and A_z are the space components, which are the three components of a 3-vector. We will sometimes write the 4-vector as

$$A = (A_t, \vec{A}) \tag{30}$$

recognizing that the last three components form an ordinary 3-vector.

Example: The difference between the components of two spacetime events forms a 4-vector

$$\Delta r = (\Delta t, \Delta x, \Delta y, \Delta z) = (\Delta t, \Delta \vec{r}) \tag{31}$$

I may sometimes refer to the four components of a 4-vector using different subscripts,

$$(A_0, A_1, A_2, A_3) \leftrightarrow (A_t, A_x, A_y, A_z) \tag{32}$$

I will try not to use this notation, but I may sometimes slip. So just remind me if I do, and remember that these are simply two different notations for describing the same object.

Not any quartet of coordinates is a 4-vector. What makes a 4-vector a 4-vector is that it behaves like the coordinate interval, (31), under a change from one inertial frame to another, that is under a Lorentz transformation. Thus if A is a 4-vector, then under a Lorentz transformation to a frame moving with speed v in the $+x$ direction, the components of A go to a new set of components A' related to the first by the Lorentz transformation

$$\begin{aligned} \Delta A'_x &= \gamma(\Delta A_x - v\Delta A_t)\,, \\ \Delta A'_t &= \gamma(\Delta A_t - v\Delta A_x)\,, \\ \Delta A'_y &= \Delta A_y\,, \quad \Delta A'_z = \Delta A_z\,. \end{aligned} \tag{33}$$

A (very) little linear algebra makes it even more obvious why this is important.

$$\begin{pmatrix} \Delta t' \\ \Delta x' \\ \Delta y' \\ \Delta z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta t \\ \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \qquad \begin{pmatrix} A'_t \\ A'_x \\ A'_y \\ A'_z \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\substack{\text{same linear} \\ \text{transformation}}} \begin{pmatrix} A_t \\ A_x \\ A_y \\ A_z \end{pmatrix} \tag{34}$$

The Lorentz transformation is described by the same matrix

$$\begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{35}$$

for every 4-vector. The same thing happens with rotations and 3-vectors:

$$\begin{pmatrix} \Delta x' \\ \Delta y' \\ \Delta z' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \qquad \begin{pmatrix} A'_x \\ A'_y \\ A'_z \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\substack{\text{same linear} \\ \text{transformation}}} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \tag{36}$$

The one "rotation matrix"

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{37}$$

describes a rotation by an angle θ about z axis for **all possible vectors!** This is at the heart of why the idea of vectors is so important.

We have now seen two things that behave like 4-vectors - the space-time interval and the energy-momentum. You will read about others in Chapter 12 of Dave Morin's book.

Now for the analog of the dot product. If we have two 4-vectors, A, from (29), and B,

$$B = (B_t, B_x, B_y, B_z) \tag{38}$$

then the combination

$$A \cdot B \equiv A_t B_t - \vec{A} \cdot \vec{B} \tag{39}$$

has the same value in any frame of reference (just as the ordinary dot product has the same value in any coordinate system). The notation here is a bit condensed. If you see a dot product between things that don't have vector indices (or bold face in David Morin's book), you should assume that the things are 4-vectors and that the dot product is the 4-dimensional invariant product defined by (39). You can check that a Lorentz transformation to a reference frame moving the in x direction does not change the value of (39). It is also obviously unchanged by rotations because the space vectors enter only through the ordinary dot product, which is unchanged by rotations. Note also that if we set $B = A$, we recover the equation for the invariant interval, s^2, and therefore we can write the invariant interval, (28), as

$$s^2 = \Delta r \cdot \Delta r \tag{40}$$

just as for 3-dimensional vectors, the square of the distance between two vectors is a dot product,

$$\ell^2 = \Delta\vec{r} \cdot \Delta\vec{r} \tag{41}$$

Right now, 4-vectors and the invariant product probably look like just a pretty mathematical analogy. But we will see shortly when we talk about using the constraints of energy and momentum conservation that they are an indispensible part of our tool box for dealing with the relativistic world.

Energy, momentum, velocity and mass

There are several ways of writing the relation between energy, momentum, velocity and mass. The one that we started with,

$$E = \frac{m}{\sqrt{1-v^2}} \qquad \vec{p} = \frac{m\vec{v}}{\sqrt{1-v^2}} \tag{42}$$

is actually not the most useful. It doesn't make sense for massless particles, such as the particles of light itself, because both the numerator and the denominators vanish as $m \to 0$. However, we can combine these into two relations that are even more general, and make sense for any m. First consider the obvious one — form the invariant product of the energy-momentum with itself. Explicit calculation gives

$$\boxed{E^2 - \vec{p}^{\,2} = m^2} \tag{43}$$

As expected, the dependence on v has gone away, because the invariant on the left hand side does not depend on the inertial frame, and thus cannot depend on how fast the particle is moving (because the speed changes when we go from one frame to another).

We can also get a relation that makes sense as $m \to 0$ by dividing the expression for momentum by the expression for energy,

$$\boxed{\vec{v} = \vec{p}/E} \tag{44}$$

These two relations are the most general formulation of the relations among energy, momentum, velocity and mass. I put boxes around them because they are very very very important. When combined with the invariant scalar product, these relations are incredibly powerful.

Massless particles

When $m = 0$, (43) and (44) are perfectly sensible, but the result is a bit odd. For $m = 0$, (43) implies that

$$|\vec{p}| = E \tag{45}$$

Then (44) implies that

$$\vec{v} = \vec{p}/|\vec{p}| = \hat{p} \tag{46}$$

which means that a massless particle always travels at the speed of light. If you think about this for a minute, it really makes your head hurt. For one thing, it means that there is no way that the state of such a particle can be specified only by its position, because massless particles with different energies move at the same speed, and thus cannot be distinguished by their speed. Thus unlike the classical particles you are used to, massless particle with different energy can have exactly the same trajectories. The resolution of this puzzle classically is to say that one shouldn't talk about massless particles at all, but just about classical waves like the electromagnetic waves you will learn more about in Physics 15b. But quantum mechanically, these particles really exist. In fact, what Einstein got the Nobel prize for was not relativity, but for his explanation of how light could knock electrons out of a metal (the photoelectric effect) because a light wave of frequency ν can be regarded of consisting of massless particles each with energy $h\nu$ where h is Planck's constant.

There is a related issue that sometimes causes confusion. Some of you have probably seen relativity before, and you may have been exposed to the rather idiotic notion of a "rest mass" that is the actual mass of the particle and a "relativistic mass" that depends on velocity. This is not useful! If you ask a physicist what the mass of the electron is, the response will be

$$m_e \approx 9.11 \times 10^{-28} \text{ g} \tag{47}$$

or

$$m_e \approx 0.511 \text{ MeV} \tag{48}$$

The response will certainly not be "Do you mean the rest mass of the electron?" or "How fast is your electron moving?"

And of course, this doesn't make any sense at all for massless particles like the photon.

Not only do these silly notions of “rest mass” and “relativistic mass” not correspond to the way physicists actually talk about these things, but the motivation for them (such as it is) is philosophically flawed. I think that the idea was to preserve the form of the equation

$$\vec{F} = m\vec{a} \tag{49}$$

for large velocities. There are two problems with this. One is that it doesn’t work, even if you allow m to depend on $\vec{v}$. You can still only preserve this form in certain special cases. But more importantly, you shouldn’t want to preserve this form anyway. We have seen that the more fundamental form that arises naturally in a Lagrangian description of mechanics is

$$\vec{F} = \frac{d}{dt}\vec{p} \tag{50}$$

We will see next week that this relation survives intact in relativistic physics. No silly definitions are required.

So if you are used to using the term “rest mass” and “relativistic mass,” you should try to get over this as soon as possible. They will only cause you grief and confusion in this course and beyond. If all else fails, perhaps hypnosis might help.

As it happens, I recently received an email from the great Russion physicist Lev Okun, asking me to comment on his paper “The mass versus relativistic and rest masses” in which he discusses the history of this issue. His argument for many years is that the notion of the “relativistic mass” $m(v)$ obscures the underlying symmetry — the 4-vector nature of $(E, \vec{p})$. In the fundamental equation

$$\boxed{E^2 - \vec{p}^{\,2} = m^2} \tag{51}$$

m is constructed using the invariant product of the 4-vector $(E, \vec{p})$ with itself. Thus the right hand side of the equation is an invariant quantity that has the same value in all reference frames. In 4-d as in 3-d, we call such a quantity a scalar. Mass is a scalar and is not conserved. Energy is a component of a 4-vector and is conserved. Don’t confuse them!

I have put Lev’s paper up with the lecture this week. If any of you have comments on the paper or suggested improvements, please let me know and I will pass them along to Lev.

lecture 13

Topics:
- Where are we now?
- Particle collisions
- K^+ decay
- Neutrino scattering
- Minimizing and maximizing
- μ decay
- Colliders

Where are we now?

Last time, we introduced the notion of 4-vectors and in particular, the 4-vector of energy and momentum that is conserved if the laws of physics are invariant under Lorentz transformations and translations in time and space. For free massive particles, these have the form

$$E = \frac{m}{\sqrt{1 - v^2}} \qquad \vec{p} = \frac{m\,\vec{v}}{\sqrt{1 - v^2}} \tag{1}$$

But I suggested that more useful and general way to think about energy and momentum of a free particle is think of the four quantities

energy	E	(2)
momentum	$\vec{p}$	
mass	m	
velocity	$\vec{v}$	

as related by the two relations

$$\boxed{E^2 - \vec{p}^{\,2} = m^2} \tag{3}$$

and

$$\boxed{\vec{v} = \vec{p}/E} \tag{4}$$

From (3) and (4), you can derive (1) if the mass is not zero. But (3) and (4) are still true even when the mass is zero and (1) doesn't make sense.

In addition, we discussed the invariant product of 4-vectors, which has the same value in all inertial frames.

Today, I now want to work through a bunch of examples of how these are useful. Most of these examples are taken from my own field of elementary particle physics, because particle physicists live with relativity every day.

Particle collisions

The relativistic energy and momentum that we derived last time are incredibly important. We derived these expressions by thinking about single free particles, but they are much more generally useful. The reason is that in almost all interesting situations, we can think of particles as free most of the time, except when they are actually colliding with one another. Then most of the time, the energy and momentum is just given by the sum of the energies and momentum of the particles. In a collision, the individual energies and momenta will change, but the total energy and momentum will be the same before and after the collision, even when new particles are created or when particles initially present are annihilated.

This is conservation of energy and momentum. Conservation means simply that when we add up the energies and momenta of the particles in the initial state of some scattering process the result is the same as if we add up the energies and momenta of the particles in the final state. The thing that I want to try to convince you of today is that it is much easier to determine the constraints that come from energy and momentum conservation if we think of the energy and momentum as a 4-vector, and use the fact that the invariant product of 4-vectors is independent of the frame.

Today, I want to do a lot of examples of the use of conservation of the energy-momentum 4-vector to analyze the decay, scattering, and production of particles. There is a very simple general idea that underlies all of these problems. The idea is to get rid of things that you don't know by using the relation $E^2 - \vec{p}^{\,2} = m^2$. Let's jump right to examples.

K^+ **decay**

There is a particle called the K^+ (pronounced "kay plus"). It is called a "strange" particle for historical reasons. This is not because it is peculiar (at least it doesn't seem peculiar any more, now that we know what it is), but because it carries a property called "strangeness". Anyway, it decays rather quickly into a pair of pions. Pions are the lightest of the particles made out of quarks and antiquarks (generically called hadrons) so they show up often. The K^+ can decay into one neutral pion (called π^0 — "pi zero"), which has a mass of about $m_{\pi^0} \approx 135$ MeV and one charged pion (called π^+ — "pi plus"), which has a mass of about $m_{\pi^+} \approx 140$ MeV. The K^+ has a mass of $m_{K^+} \approx 494$ MeV. Now suppose that the decay of the K^+ occurs at rest. What are the energies of the two pions? This is a typical sort of question in what might be called decay kinematics. To answer such questions, we think about 4-vectors and use conservation of energy and momentum. Let us call the energy-momentum 4-vectors K for the K^+ and π_+ and π_0 for the π^+ and π^0 respectively. Conservation of energy and momentum is the statement that the 4-dimensional vectors satisfy

$$K = \pi_+ + \pi_0 \tag{5}$$

This is a short-hand for four equations, conservation of energy and conservation of each of the three components of momentum. Note that it is true in any frame of reference. In the rest frame, the 4-vectors look like

$$
\begin{aligned}
K &= (m_K, 0) \\
\pi_+ &= (E_+, \vec{p}_+) \\
\pi_0 &= (E_0, \vec{p}_0)
\end{aligned}
\tag{6}
$$

We have used a standard trick here — one that you should be familiar with from our rules of coherence. If you don't know something, give it a name! Now (5) can be used to say things about (6), for example, $\vec{p}_+ = -\vec{p}_0$. But let us try to resist this temptation. In problems like this, it is often convenient to manipulate the 4-vectors symbolically for a while before actually doing the calculation. The idea of such manipulations is to be able to use the invariant product to compute what you want to know without having to calculate what we don't care about. Here for example, suppose that we first want to calculate the energy of the π^+, which we have called E_+. If we could calculate the value of the invariant product $K \cdot \pi_+ = m_K\, E_+$, that would immediately give us E_+. So suppose that we rewrite (5) as

$$
K - \pi_+ = \pi_0 \tag{7}
$$

Now if we take the invariant product of each side of this equality with itself, we will get terms involving $K \cdot \pi_+ = m_K\, E_+$:

$$
\begin{aligned}
(K - \pi_+) \cdot (K - \pi_+) &= K \cdot K - 2\, K \cdot \pi_+ + \pi_+ \cdot \pi_+ \\
&= m_K^2 - 2\, K \cdot \pi_+ + m_{\pi^+}^2 = \pi_0 \cdot \pi_0 = m_{\pi^0}^2
\end{aligned}
\tag{8}
$$

Now we can solve this for $K \cdot \pi_+$

$$
K \cdot \pi_+ = \frac{m_K^2 + m_{\pi^+}^2 - m_{\pi^0}^2}{2} \tag{9}
$$

or

$$
E_+ = \frac{m_K^2 + m_{\pi^+}^2 - m_{\pi^0}^2}{2 m_K} \tag{10}
$$

Easy, no? Now we can get E_0 either by repeating the same calculation with + and 0 interchanged, or by using energy conservation. The result is

$$
E_0 = \frac{m_K^2 + m_{\pi^0}^2 - m_{\pi^+}^2}{2 m_K} \tag{11}
$$

If you were asked to do so, you could now go on and calculate the magnitudes of the momenta of the pions by using $E^2 - \vec{p}^{\,2} = m^2$. From there, you could calculate the speeds, although this is seldom very interesting in such collisions. You cannot calculate the direction of the momentum or velocity, because this is actually quite random. The K^+ is a particle with no angular momentum. When it is at rest, there is no vector associated with it. And therefore there is no direction picked out for its decay products. They go off at random with equal probability in all directions. That is quantum mechanics. God plays dice with the universe.

Neutrino scattering

Here is an example of an interesting scattering process. Neutrinos are very light particles. Until recently, we thought that they might be massless, like photons. But it now appears that the neutrinos have tiny masses. Furthermore, these masses are very peculiar. There are three different kinds of neutrinos: ν_e (an “electron neutrino”); ν_μ (a “mu neutrino”); and ν_τ (a “tau neutrino”). The names refer to the processes in which these neutrinos are produced, which involve respectively the electron and the heavier versions of the electron, the μ (“mu”) and the τ (“tau”). The tiny masses do not respect these distinctions, and they produce bizarre quantum mechanical mixing between these different types of neutrinos. But if it were not for these weird quantum mechanical effects, we could ignore the neutrino masses altogether. So that is what we will do in this course. We will simply pretend that neutrinos are massless, which is an excellent approximation for the sort of questions that we can ask and answer in this course.

In spite of the fact that neutrinos have no electric charge and almost no mass, it is possible to make beams of neutrinos. In fact, one of my colleagues, Gary Feldman, is part of a large project that involves making a neutrino beam at Fermilab outside Chicago and aiming it at a large detector in an underground mine in northern Minnesota. This should tell us more about the peculiar neutrino masses I mentioned earlier.

But here, I just want to note that with these beams, we can observe bizarre processes such as the scattering of a ν_μ with energy E from an electron at rest to produce a final state consisting of a ν_e and a μ. The μ has a mass m_μ about 207 times the mass of the electron, m_e.

Now, a question that you might ask about this process is the following. Suppose that you see a ν_e in the final state flying off at an angle ϕ from the initial direction of the ν_μ: as shown:

Initial state final state (12)

What does energy and momentum conservation tell you about the energy of the ν_e in the final state?

The 4-vectors look like

$$\begin{aligned} e &= (m_e, 0, 0, 0) \\ \nu_\mu &= (E, E, 0, 0) \\ \nu_e &= (E_1, E_1 \cos\phi, E_1 \sin\phi, 0) \\ \mu &= (E_2, p_2 \cos\theta, -p_2 \sin\theta, 0) \end{aligned} \tag{13}$$

Some comments about this are in order. I have chosen to put the initial ν_μ momentum along the x axis. That is no problem, because I can rotate my coordinate system to make it so. Likewise, I have assumed that the scattering takes place in the x-y plane, which I can again do by just rotating the coordinate system. I have put in the information that the neutrinos are massless by taking the lengths of their momentum vectors to equal their energies. If I have not done this and just given the momenta names, we would have quickly gotten to this point when we imposed $E^2 - \vec{p}^{\,2} = m^2 = 0$ on these 4-momenta. I have also, in my head, imposed a little bit of energy momentum conservation by writing μ in the x-y plane (although we won't use this at all). Again, if we had put in a z component for μ, we would have quickly realized that it must be zero because all the other 4-vectors have zero z component by construction.

Now, we could simply impose energy and momentum conservation of the rest of (13), and try to solve the equations. But it is better to think. For example, we can easily find E_1, because energy-momentum conservation

$$\nu_\mu + e = \mu + \nu_e \tag{14}$$

implies

$$\nu_\mu + e - \nu_e = \mu \tag{15}$$

This is a good way to write things, because when we take the invariant product of each side with itself, all the nonsense in μ (which we don't know yet), drops out, and we can write

$$(\nu_\mu + e - \nu_e) \cdot (\nu_\mu + e - \nu_e) = \mu \cdot \mu = m_\mu^2 \tag{16}$$

The left hand side of (16) is

$$\nu_\mu \cdot \nu_\mu + e \cdot e + \nu_e \cdot \nu_e + 2\,\nu_\mu \cdot e - 2\,\nu_\mu \cdot \nu_e - 2\,e \cdot \nu_e \tag{17}$$

which with (16) implies

$$m_e^2 + 2m_e E - 2E\,E_1(1 - \cos\phi) - 2m_e E_1 = m_\mu^2 \tag{18}$$

so that

$$E_1 = \frac{m_e^2 + 2m_e E - m_\mu^2}{2E\,(1 - \cos\phi) - 2m_e} \tag{19}$$

Practically speaking, this is a bit of a swindle. There is nothing wrong with the calculation above, but a particle physicist would never ask you to calculate things in terms of the angle of the final state neutrino. This is because neutrinos are very hard to see. They very seldom interact with anything. That is how they will manage to get from Chicago to northern Minnesota. But there are many ways of making the direction of charged particle track show up, all making use of the electric charge and its interactions. So it would make more sense physically to ask you to find things in terms of the μ angle, θ, or its energy, E_2, or momentum, p_2. This is a little more involved algebraically, but the principle is the same.

Minimizing and maximizing

There is an interesting and useful class of questions in which the kinematics does not completely fix the interesting quantities, and you have to think about how to make them bigger or smaller. Here is a typical situation. Suppose that a particle with mass m_1 and energy E collides with a particle with mass m_2 and the collision produces a final state with n particles with masses μ_1, μ_2, $\ldots$ μ_n. What is the minimum energy E required for a process that produces these particles in the final state to take place? And what is the total energy of the process in the zero momentum frame?

First let me write down the answer, which is rather simple. The total energy of the process in the zero momentum frame is at least

$$E_{\vec{p}=0} = \sum_{j=1}^{n} \mu_j \equiv M_{\rm min} \tag{20}$$

and the energy E of the particle with mass m_1 in the frame in which particle 2 is as rest is

$$\frac{M_{\rm min}^2 - m_1^2 - m_2^2}{2m_2} \tag{21}$$

Let us begin the proof by discussing the question in general. The key to problems like this is to treat the whole final state as a single entity. The final state, whatever it is, will have some total energy-momentum 4-vector

$$T = (E_T, \vec{P}_T) = \sum_{j=1}^{n} (E_j, \vec{p}_j) \tag{22}$$

Then we can define a "total mass" M_T for the final state using the fundamental relation between energy, momentum and mass —

$$E_T^2 - \vec{P}_T^{\,2} = M_T^2 \tag{23}$$

Then in terms of M_T, this problem is formally equivalent to the problem of producing a particle with mass M_T by colliding a particle with mass m_1 and energy E with a mass m_2 particle at rest. Define the 4-momenta as

$$P_1 = (E, p) \qquad P_2 = (m_2, 0) \tag{24}$$

Then 4-momentum conservation implies

$$P_1 + P_2 = T \tag{25}$$

Taking the invariant product of each side gives

$$P_1 \cdot P_1 + 2P_1 \cdot P_2 + P_2 \cdot P_2 = m_1^2 + 2m_2\, E + m_2^2 = M_T^2 \tag{26}$$

$$E = \frac{M_T^2 - m_1^2 - m_2^2}{2m_2} \tag{27}$$

which should remind you of (21). Evidently, to minimize E we need to minimize M_T, and if we find that the minimum value of M_T is the $M_{\rm min}$ in (20), then we are done. So how do we do that?

The advantage of considering M_T is that it is an invariant quantity. It has the same value in all frames of reference. In particular, it looks the same in the zero momentum frame. Thus in this frame, the total energy is just M_T

$$T' = (E'_T, \vec{P}'_T) = \sum_{j=1}^{n} (E'_j, \vec{p}'_j)' = (E'_T, 0) = (M_T, 0) \tag{28}$$

But (28) implies that

$$M_T = \sum_{j=1}^{n} E'_j \tag{29}$$

Then since the energy of a particle is always greater than or equal to its mass, (29) implies

$$M_T \geq \sum_{j=1}^{n} \mu_j = M_{\rm min} \tag{30}$$

If the particles in the final state are all massive — that is if they have $\mu_j > 0$ for all j, then the bound (30) is saturated when all the particles are at rest in the zero momentum frame and this is the best we can do in minimizing M_T.

If there are massless particles, we cannot take their momenta to be zero, but we can take them to be very very small. In this way, we can come arbitrarily close to the theoretical minimum. If we now boost this final state by a Lorentz transformation, all of the massive particles will be traveling with the same velocity, and the massless particles will still have arbitrarily small energy and momentum. This is the best we can do.

Here is a simple (and classic) example. Suppose that you want to make antiprotons. You can do this by hitting a proton at rest with a proton of energy E to produce a final state consisting of three protons, each with mass m_p, and one antiproton (also with mass m_p). What is the minimum energy required to make antiprotons this way? I should say that this is actually the way that antiprotons are made at Fermilab and CERN (where for example Professor Gabrielse uses them to make antihydrogen and study its properties). Because all of the final state particles are massive, we can take all four particles to be moving with the same velocity, v, and the total mass will then be $M_T = 4m_p$. Putting this into (27) gives the result:

$$E = \frac{M_T^2 - 2m_p^2}{2m_p} = \frac{16m_p^2 - 2m_p^2}{2m_p} = 7m_p \tag{31}$$

Note we have obtained a little bit more than the answer we were looking for. We actually know something about the final state of the four particles when (31) is satisfied — that all the particles are moving with the same velocity.

To think about this another way, imagine slowly cranking up the energy of our proton beam until we start to produce antiprotons. When this first happens, at an energy given by (31), the four particles have the same velocities, but as we go to higher energy, the velocities are the same and some of the energy of our original beam is wasted in producing the extra energy of that relative motion, rather than just going into making antiprotons.

This result is sort of neat, but the point is the style of argument. The idea is to think of whole collections of particles as having a "mass" — computed from the total energy momentum 4-vector.

μ decay

We have already talked about the fact that the heavy version of the electron called the muon, μ, is unstable. It decays into an electron, a muon neutrino, and an electron antineutrino:

$$\mu^- \to e^- + \nu_\mu + \bar{\nu}_e \tag{32}$$

This process is actually very closely related to the inverse of the scattering process we discussed earlier,

$$\nu_e + \mu^- \to \nu_\mu + e^- \tag{33}$$

There is a sense in which the $\bar{\nu}_e$ in (32) is related to a ν_e traveling backward in time. That is to say that a $\bar{\nu}_e$ in the final state is related to a ν_e in the initial state. If we made this change in (32), we would get just the inverse of the process (33). In some sense, this is why antiparticles must exist for every type of particle. This is actually related to a discrete symmetry called CPT which stands for

Charge Conjugation-**P**arity-**T**ime reversal

which might be an exact symmetry of the world.

At any rate, we can ask the usual sort of questions about this. Here is a modest list. Suppose that the μ decays while it is at rest.

1. What is the minimum possible energy of the electron?

2. What is the maximum possible energy of the electron?

3. What is the minimum possible energy of one of the neutrinos? I really mean neutrino or antineutrino here, but it takes too long to say that each time, and they are both nearly massless anyway, so I won't bother.

4. What is the maximum possible energy of one of the neutrinos?

5. Consider the total energy and momentum of the two neutrinos as a 4-vector. What is the maximum possible value of the invariant product of this 4-vector with itself, $E^2 - \vec{p}^{\,2}$? Or equivalently, what is the maximum "mass" of the two neutrino system?

6. What is the minimum "mass" of the two neutrino system?

The idea of all these can be found by thinking about the analysis we just did for the mass of a system of particles. Let's take them one at a time.

1. **What is the minimum possible energy of the electron?** The smallest energy the electron could possibly have is m_e, which it would have if it were at rest. Is this possible? Sure! We can conserve energy and momentum if all the rest of the energy goes into two back-to-back neutrinos, so the 4-momenta would look like

$$\mu = (m_\mu, 0) \qquad e = (m_e, 0) \qquad \nu_\mu = (E, E\hat{v}) \qquad \bar{\nu}_e = (E, -E\hat{v}) \tag{34}$$

   Conservation of energy and momentum works if $E = (m_\mu - m_e)/2$.

2. **What is the maximum possible energy of the electron?** To get the maximum possible energy for the electron, what we want is that the effective mass of the rest of the stuff in the decay, the two neutrinos, should be as small as possible. But if the two neutrinos have parallel momenta, the mass of the two-neutrino system is zero. That is the best we can do. The process then looks like

$$\mu = (m_\mu, 0) \qquad e = (E, p\hat{v}) \qquad \nu_\mu = (xp, -xp\hat{v}) \qquad \bar{\nu}_e = (yp, -yp\hat{v}) \tag{35}$$

   where $x + y = 1$. We can calculate E easily as we have done in other problems if we note that $\mu - e$ has mass 0, so that

$$m_\mu^2 - 2m_\mu E + m_e^2 = 0 \quad \Rightarrow \quad E = \frac{m_\mu^2 + m_e^2}{2m_\mu} \tag{36}$$

3. **What is the minimum possible energy of one of the neutrinos?** Either of the neutrinos can have arbitrarily small energy and momentum. There is enough freedom to satisfy energy and momentum conservation with the other two carrying the load.

4. **What is the maximum possible energy of one of the neutrinos?** This is actually related to the previous question. The maximum neutrino energy arises when the neutrino recoils against the minimum possible mass, which is the electron mass, with the other neutrino carrying negligible energy and momentum:

$$\mu = (m_\mu, 0) \qquad e = (E, p\hat{v}) \qquad \nu_\mu = (p, -p\hat{v}) \qquad \bar{\nu}_e = (\approx 0, \approx 0) \tag{37}$$

   Now the mass of the 4-vector $\mu - \nu_\mu$ is μ_e, so

$$m_\mu^2 - 2m_\mu p = m_e^2 \quad \Rightarrow \quad p = \frac{m_\mu^2 - m_e^2}{2m_\mu} \tag{38}$$

5. **What is the maximum “mass” of the two neutrino system?** We have really done this one already. The two-neutrino system has its maximum mass when it and the electron are both at rest, and the mass of the two-neutrino system is $m_\mu - m_e$.

6. **What is the minimum “mass” of the two neutrino system?** We’ve done this one also. If the two neutrino momenta are parallel, the mass is zero

Colliders

Last time, we looked at the process

$$J/\psi \to e^- + e^+ \tag{39}$$

where e^- is an electron, e^+ is a positron, the antiparticle of the electron, and J/ψ is a particle that has two names for historical reasons. It was discovered allegedly independently at Brookhaven and at SLAC. Anyway, it has a mass of $m_{J/\psi} \approx 3097$ MeV.

$$e^- + e^+ \to J/\psi \tag{40}$$

Now suppose that we let positrons with energy E collide with electrons at rest to produce J/ψs. What energy is required? The 4-vectors are as follows:

$$e^- = (m_e, 0) \qquad e^+ = (E, \vec{p}) \qquad J/\psi = (E', \vec{p}\,') \tag{41}$$

Now we can use energy-momentum conservation

$$e^- + e^+ = J/\psi \tag{42}$$

Taking the invariant product of each side with itself gives

$$m_e^2 + 2m_eE + m_e^2 = m_{J/\psi}^2 \tag{43}$$

so that

$$E = \frac{m_{J/\psi}^2 - 2m_e^2}{2m_e} \approx 9385 \text{ GeV} \tag{44}$$

This is a huge energy scale, beyond what is presently available at accelerators. The problem is that the electron is very light. A collision between a very high energy electron and an electron at rest is a bit like a nonrelativistic collision between a moving truck and a feather. Very little energy is actually transfered in such a collision.

However, it is much easier to produce the J/ψ in a collider, in which an electron and positron collide with equal and opposite velocities and momenta. In this case, the 4-vectors look like

$$e^- = (E, \vec{p}) \qquad e^+ = (E, -\vec{p}) \qquad J/\psi = (m_{J/\psi}, 0) \tag{45}$$

Here, energy-momentum conservation implies

$$E = m_{J/\psi}/2 \approx 1.55 \text{ GeV} \tag{46}$$

Colliding beams, in this case, make a huge difference. The difference is that in the collision with a particle at rest, much of the energy of the incoming particle is wasted producing kinetic energy of the collision products. This effect exists in Newtonian physics also, but it is much worse at high energies where relativistic physics takes over. Note that the problem is particularly bad for the electron, because it is the lightest particle with electric charge. As (44) shows, the energy required to produce a heavy particle of mass M in a fixed target collision between particles of mass m, much less than M is inversely proportional to m, so the lightness of the electron is a terrible problem. But if you want to have the capability to produce the heaviest possible things, colliding beams are essential no matter what you are colliding.

Why is this interesting? What are these heavy particles that particle physicists make, and why would you want to make them? The first thing to say is that we don't need heavy particles to make heavy things. The things in the universe that are much heavier than protons and neutrons and electrons - from molecules to galaxies - are made by putting many copies of these light things together. This can be done in a practically infinite number of way and produces lots of interesting physics. But the heavy particles of particle physics are very different. They are not just light things put together in interesting ways. And there are only a very few of them, with completely well defined masses and properties. To make one of these particles with mass M, we must not only have enough energy, M, in the zero momentum frame, but we must also concentrate that energy in a tiny region of space and time, of size $1/M$ (in particle physics units). They are genuinely new indivisable degrees of freedom that appear only when look at the world at large energies and small distances. They are not useful, unless you want to know how the world works.

lecture 14

Topics:

- Where are we now?
- Causality and rigid bodies
- Forces
- Relativistic strings
- Color force
- A relativistic oscillator
- The moving oscillator

Where are we now?

So far we have been discussing mostly the kinematics of special relativity. This week we will talk a bit about dynamics. I will talk about forces and the work-energy theorem in relativity in general. Then I will introduce the rather bizarre idea of relativistic strings. I have several reasons for wanting to do this. The first is that relativistic strings will allow us to do a completely honest analysis of a relativity "paradox" similar to some you have read about in David Morin's notes. I have avoided these until now. Most of them seem more paradoxical than they really are because we are used to thinking of large objects as rigid. This makes no sense in relativity, and we will start today by reviewing that argument. But with relativistic strings, we will be able to build objects that get large, but can still be analyzed in a way completely consistent with special relativity. The second reason for discussing relativistic strings is that they are an incredibly hot topic in theoretical physics and mathematics these days. Needless to say, we won't get very far along this path. But I thought that you might enjoy seeing just a tiny bit of it. Believe it or not, the third reason for introducing relativistic strings is that if you let youself go and really imagine that this stuff exists, it will actually help you think about forces in relativity. In particular you can use it as a mnemonic to help you remember how to Lorentz transform force.

Causality and rigid bodies

Some of our intuition about the way things work depends on the rigidity of solids. This leads to some of the paradoxes in relativity, because the notion of a rigid body makes absolutely no sense in relativity. This is a fun issue because it is related to the philosophical issue of causality — cause and effect — and is closely tied to the fact that information cannot travel faster than light.[1] The point is this. Two events that are separated by a space-like interval cannot have any effect on one another if we accept the principal that effects must come after their causes. For any two space-like separated events, there exist frames of reference in which either event comes first. We have already seen that given two space-like separated events, there exists a frame in which the two events occur at the same time. From this frame, we can look at frames which are moving in the direction of the

[1] We will return to this important principle several time in various contexts.

space separation between the two events. Then in a frame moving towards one event, that event must come first. You can see this directly from the Lorentz transform — it is related to the minus sign in (for a separation in the z direction)

$$\Delta t' = \frac{1}{\sqrt{1-v^2}}(\Delta t - v\,\Delta z) \tag{1}$$

Or we can see it qualitatively from the classic argument for relativity of simultanaeity in which we send light pulses from the center to the two events.

Now the problem with the idea of a rigid body is this. If you push on one end of a rigid body, the whole body starts to accelerate. This is impossible, because the two events that mark the beginning of acceleration at the two ends of the rigid body are separated by a space-like interval. In some frames, the beginning of acceleration of the far end of the rigid body comes before the beginning of acceleration where you push. So your push cannot cause the acceleration at the far end. And these are physical events. Both the push that causes the acceleration and the acceleration at the other end involve non-zero forces that are not just artifacts of a particular inertial frame. Thus rigid bodies are impossible unless effects can happen before their causes — which is not a good idea.

This is closely related to the fact that information cannot travel faster than light. We have just argued that an effect and its cause cannot be separated by a space-like interval. They must therefore be separated by light-like or time-like interval. If the interval is time-like, a massive particle can travel at a velocity less than the speed of light from the spacetime coordinates of the effect (the "effect event" if you like) and the spacetime coordinates of the cause (the "cause event"), because these two events can be two ticks of the same clock. If the interval is light-like, light can travel from the cause event to the effect event. But these are the only possibilities. If the information travels from the cause event to the effect event at a speed greater than the speed of light, that means that the distance L between the cause and the effect is greater than c times the time difference T between the two events, which means that $c^2T^2 - L^2 < 0$ and the interval is spacelike. Thus the far end of the body cannot possibly know that the body has been pushed any sooner than it takes light to travel from the push to the other end. Otherwise there would be a frame in which the effect happens before the cause.

What is really going on here is this. Even in non-relativistic mechanics, the idea of a rigid body is an idealization. No real body is perfectly rigid. In fact, when you push on one end of a real object, compression waves travel through the object to transmit the information to all the parts. These waves travel at a characteristic speed (like "sound waves") that depends on how heavy and how stiff the body's material is. The stiffness is like the spring constant of a spring, and the stiffer the material, the faster the waves. But in non-relativistic physics, there is no reason why you cannot imagine a material that is arbitrarily stiff, so that the speed of these waves goes to infinity and in the limit, the body becomes truly rigid. In relativity, this is impossible, because these waves carry information, and thus cannot travel faster than light.[2] So the most accurate statement is that

[2] You will learn in Physics 15c, if you have not already, that it is really a bit complicated. What cannot be faster than light is actually the group velocity, which is not necessarily the same as the velocity at which the wave crests move, which is called the phase velocity. But it is the group velocity that describes how fast information can be transmitted,

relativity makes it impossible to build arbitrarily stiff material. Even the forces that hold matter together are built on the principle that information cannot travel faster than light.

The notion that information cannot travel faster than light will be very important when we talk about the early universe, which I plan to do at the end of the course.

Forces

As we discussed when we talked about the Lagrangian for a massive relativistic particle, we can add additional terms that describe forces, and then as usual, the force on the particle will be the rate of change of the momentum. It is important to note that because the energy and momentum in special relativity are derivable from a Lagrangian, we can use the same relation between force on a particle, the particle's momentum and the power transmitted to it by the force in relativistic physics that we do in nonrelativistic physics,

$$\vec{F} = \frac{d\vec{p}}{dt} \tag{2}$$

$$\vec{v} \cdot \vec{F} = \frac{dE}{dt} \tag{3}$$

The first, (2), can be regarded as a definition of what we mean by force. It is just a rewriting of the Euler-Lagrange equation. The second, (3), which says that the dot product of the force and the velocity is the power supplied to the system, then follows from the definition of E. It is also consistent with the fundamental relations between energy, momentum, velocity and mass,

$$E^2 = p^2 + m^2 \qquad \vec{v} = \vec{p}/E \tag{4}$$

We can see this consistency explicitly if we differentiate both sides of (4) with respect to t, to get

$$E\,\frac{dE}{dt} = \vec{p} \cdot \frac{d\vec{p}}{dt} = \vec{p} \cdot \vec{F} \tag{5}$$

or

$$\frac{dE}{dt} = \frac{\vec{p}}{E} \cdot \vec{F} = \vec{v} \cdot \vec{F} \tag{6}$$

This last can be rewritten by multiplying by dt as the relation between work done by the force and the change in energy.

$$dE = d\vec{r} \cdot \vec{F} \tag{7}$$

Again, this is not surprising. It had to work because the energy and momentum of the relativistic particle are derivable from a Lagrangian.

Let's discuss an example of all this. Suppose that a particle of mass m is subject to a constant force F_0 in the x direction. From (2) we can immediately conclude that

$$\vec{p} = \hat{x}\,F_0\,t \tag{8}$$

which is what matters in this argument.

Now suppose that we want to figure out the distance, $x(t)$, the particle has traveled. We can do this by using (4) and (7).

$$\begin{aligned} F_0\, x(t) &= \int_0^t dt'\, \frac{dE(t')}{dt'} = E(t) - E(0) \\ &= \sqrt{p^2 + m^2} - m = \sqrt{F_0^2\, t^2 + m^2} - m \end{aligned} \tag{9}$$

Thus

$$x(t) = \frac{\sqrt{F_0^2\, t^2 + m^2} - m}{F_0} \tag{10}$$

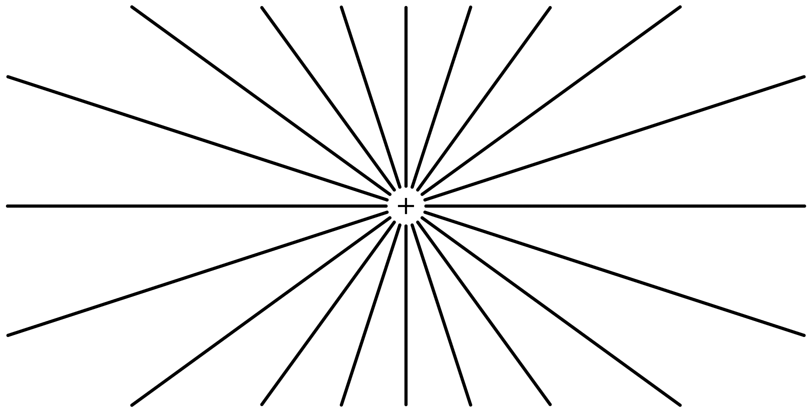
This kind of trick with the work-energy relation is often useful. Study it. On your own, as practice with the Taylor expansion, you should verify that (10) reduces to the expected non-relativistic result in the appropriate limit.

Relativistic strings

When you pull on an elastic material object, it stretches, and stores the energy you feed into it in potential energy associated with the quantum mechanical electromagnetic interactions that hold matter together. In relativistic physics, one can imagine a different scenario. Imagine stuff that when you pull on it, creates more of itself. In this case, the energy is stored in the mass energy of the new stuff that is created. This sounds absolutely crazy. But there are good reasons to believe that this stuff really exists. It is called "relativistic string". In particular, the glue that holds quarks and antiquarks together seems to behave like a relativistic string. Let me discuss just a bit how this goes.

The simplest state involving quarks is a state of one quark and one antiquark. In many ways, this is like a bound state of an electron and its antiparticle, the positron, which is called positronium. The force that holds quarks and antiquarks together not the same force that holds the electron and the positron together in positronium. The force in positronium is just the Coulomb force that you are familiar with between ordinary charged particles. The force between quark and antiquark is much stronger and has other peculiar properties, as we will see. Nevertheless, both forces can be described in terms of fields and "field lines". Let me warm up by reminding you about our discussion of field lines for the ordinary electric force.

Field lines begin at positive charges and end at negative charges, but otherwise are continuous. The direction of the field lines at any point in space indicates the direction of the electric field at that point. And the electric field is the force per unit charge on a charge placed at the point in question. The density of the field lines indicates the magnitude of the electric field. So, for example, a single positively charged particle at rest has field lines spreading out from the positive charge in all directions, uniformly. The density of the fields lines drops off like $1/r^2$ because the field lines at a distance r from the charge are spread out over a sphere of radius r. Thus the electric field drops off like $1/r^2$, which in turn is the reason for the $1/r^2$ behavior of the Coulomb force. Notice that this depends on being in 3 space dimensions. If I draw a two dimensional representation of the field lines, it will give the general idea, but will be quantitatively wrong because the lines only spread out over the surface of a circle as they go out. Nevertheless, these two dimensional pictures can help us to visualize what is happening with the field line, like so

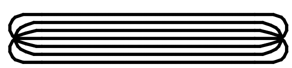


(11)

In positronium, it is the electric force associated with these field lines that holds the system together. The electron and positron have opposite charges, so the force is attractive. When the electron and the positron are pulled apart, the force decreases like $1/r^2$.

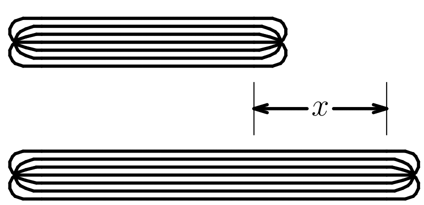
Color force

In the quark-antiquark system, the force is in some ways similar to the electric force — we call it a "color" force. The term "color" is rather whimsical, having nothing to do with real colors.[3] Very close to the quark and antiquark, there is a color field that spreads out from the charges like the electric field from a positron or electron. But for reasons that are complicated and only partly understood, at distances far from the quark and antiquark, the color electric field lines organize themselves into a tube, as shown below:



(12)

This tube behaves approximately like a relativistic string (the difference is just the end effects near the quark and antiquark, and we will always ignore these). When I pull the quarks farther apart, the tube just gets longer. The energy is stored in the color electric field of the string, but that just gives some fixed energy per unit length because the tube looks the same everywhere except at its ends, as illustrated below:



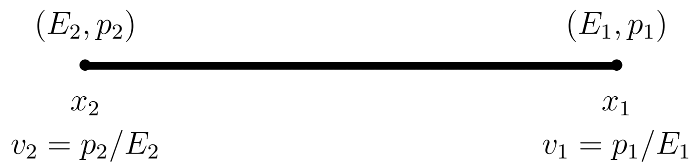
(13)

The extra energy in the stretched string is xT where T is the energy per unit length of the string. Because the energy of the string increases linearly with length, that means that the string exerts a

[3]It is actually based on an analogy with the way the primary colors can be combined to produce "colorless" white light, but this won't be important for what I am going to tell you.

constant force T at its ends. Thus the energy per unit length of the string in its rest frame, T, is called the “string tension”.

Relativistic strings are extremely complicated if you think about how they bend and oscillate. But it is not too hard to understand how they move if they are stretched in a single dimension, and if they have particles (like the quarks) stuck to their ends. This is the only thing we will do in this course. In this case, we can just apply what we know about force and rate of change of momentum to understand what is going on. The situation is easiest to understand if the particles are only moving in the dimension in which the string is stretched, so the string has no transverse motion. The system then looks like this (let’s assume that everything is in the x direction and from here on, we will forget about any width to the string and just draw it as a fat line):



(14)

where the momenta and velocities all refer only to the x component. Note that each of the particles may be moving in either the plus or minus x direction.

Relativistic strings have an absolutely bizarre property. You cannot tell whether they are moving in the direction in which they are stretched. That is because they are completely featureless. There is nothing to show you that they are moving. You may be able to see this for the string we discussed above made out of a tube of field lines. If the system is moving, the field lines are contracted, but as long as there is no transverse motion, the lines don’t get closer together, and therefore the density doesn’t change. But it is the density that determines the strength of the field, so that means that the field doesn’t change and the energy per unit length is the same.

So in (14), you need not think of the string as moving where it is attached to the particles at the ends. Rather, you can just as well think of the string as just sitting there and getting created or eaten up as the particles move. It doesn’t matter which picture you use. These two very different sounding picture are actually physically equivalent!

Because the string is being created and eaten up as the particles move, there is a force on each of the particles equal to the string tension, T, which is the mass per unit length of the string. Thus

$$\frac{dp_1}{dt} = -T \qquad \frac{dp_2}{dt} = T \tag{15}$$

where the signs work this way because the string is always trying to contract into nothingness to reduce its energy, and

$$\frac{dE_1}{dt} = -T\, v_1 \qquad \frac{dE_2}{dt} = T\, v_2 \tag{16}$$

or equivalently

$$\frac{dE_1}{dx_1} = -T \qquad \frac{dE_2}{dx_2} = T \tag{17}$$

Note that the total momentum is conserved, and the total energy is conserved if the energy stored in the string is properly included. The energy change in (17) just comes from the string that is eaten up or created.

A relativistic oscillator

I will now give you what I find a very amusing puzzle. Imagine that I have piece of relativistic string with string tension T stretched along the x axis, and that stuck to the two ends, there are massless particles, a quark, q, and an antiquark, $\bar{q}$, which also move along the x axis. If the total momentum of the quark and antiquark is zero, this is a relativistic oscillator. Suppose that at time $t = 0$ (in the lab frame), the quark and antiquark are both at $x = 0$, with the quark moving in the $+x$ direction with momentum κ and the antiquark moving in the $-x$ direction with momentum $-\kappa$. The corresponding energies are both κ, because these are massless particles. At this point, there is no string, because both particles are at the origin, so the total energy and momentum of this system is

$$(\kappa, \kappa\hat{x}) + (\kappa, -\kappa\hat{x}) = (2\kappa, 0) \tag{18}$$

Now as the two particles fly outward at the speed of light, they leave string in their wake. As they move they lose energy T per unit length because the string does work on the particles. The magnitude of the momentum is always equal to the energy because the particles travel at the speed of light. Thus the particles travel at the speed of light in the same direction until at $t = \kappa/T$, having traveled a distance κ/T, they have lost all their energy and momentum, and they turn around and start back in towards the origin at the speed of light. At the turn-around point, all their energy is stored in the string, and the quark and antiquark carry negligible energy and momentum. At $t = 2\kappa/T$, the two particles are back at the origin with the quark now moving in the $-x$ direction and the antiquark in the $+x$ direction. We assume that they can pass right through each other and continue on until at $t = 3\kappa/T$, they turn around again. At $t = 4\kappa/T$, the particles are back where they started at, at $x = 0$, and the whole process repeats. Thus this is an oscillator with period $4\kappa/T$. This is summarized in the table below, where we describe these things in terms of the relevant space-time events

events	q	$\bar{q}$
together at $t = 0$, $x = 0$	$(0, 0)$	$(0, 0)$
turn around	$\left(\frac{\kappa}{T}, \frac{\kappa}{T}\right)$	$\left(\frac{\kappa}{T}, -\frac{\kappa}{T}\right)$
at $x = 0$ backwards	$\left(2\frac{\kappa}{T}, 0\right)$	$\left(2\frac{\kappa}{T}, 0\right)$
turn around again	$\left(3\frac{\kappa}{T}, -\frac{\kappa}{T}\right)$	$\left(3\frac{\kappa}{T}, \frac{\kappa}{T}\right)$
back at 0	$\left(4\frac{\kappa}{T}, 0\right)$	$\left(4\frac{\kappa}{T}, 0\right)$

(19)

This object is animated in STRINGY.

The moving oscillator

This is a wonderful system to think about, because even though it gets big, we can understand exactly how it looks in different frames, because we understand all the forces that hold the system together. But here is the puzzle. How can this thing possibly move? Massless particles always travel at the speed of light in ANY frame. When we go to another frame, for example one moving with velocity v in the $+x$ direction, in which we are moving towards the oscillator, we should see it moving with velocity $-v$. But at $t = 0$, the two quarks are both moving out from the origin at the speed of light, in both frames, so the oscillator appears to be standing still! There is more string growing on both sides - but symmetrically, so that the center of mass remains fixed at the origin. What is going on?????

As with many relativity puzzles, the resolution lies in the relativity of simultaneity. Let us see how this works by explicitly doing the Lorentz transformations of the events in (19) to a frame moving with velocity v in the $+x$ direction. The first thing you notice is that (19) is a little misleading, because some of the rows refer to single events, when the quark and antiquark are in the same place at the origin, but others actually refer to two events, when the quark and antiquark turn around on opposite sides of the system. With this confusion corrected, the result looks like this:

lab frame		moving frame	
q	$\bar{q}$	q	$\bar{q}$
$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$
$\left(\frac{\kappa}{T},\frac{\kappa}{T}\right)$		$\gamma\Big((1-v)\frac{\kappa}{T},-(v-1)\frac{\kappa}{T}\Big)$	
	$\left(\frac{\kappa}{T},-\frac{\kappa}{T}\right)$		$\gamma\Big((1+v)\frac{\kappa}{T},-(v+1)\frac{\kappa}{T}\Big)$
$\left(2\frac{\kappa}{T},0\right)$	$\left(2\frac{\kappa}{T},0\right)$	$\gamma\left(2\frac{\kappa}{T},-2v\frac{\kappa}{T}\right)$	$\gamma\left(2\frac{\kappa}{T},-2v\frac{\kappa}{T}\right)$
	$\left(3\frac{\kappa}{T},\frac{\kappa}{T}\right)$		$\gamma\Big((3-v)\frac{\kappa}{T},-(3v-1)\frac{\kappa}{T}\Big)$
$\left(3\frac{\kappa}{T},-\frac{\kappa}{T}\right)$		$\gamma\Big((3+v)\frac{\kappa}{T},-(3v+1)\frac{\kappa}{T}\Big)$	
$\left(4\frac{\kappa}{T},0\right)$	$\left(4\frac{\kappa}{T},0\right)$	$\gamma\left(4\frac{\kappa}{T},-4v\frac{\kappa}{T}\right)$	$\gamma\left(4\frac{\kappa}{T},-4v\frac{\kappa}{T}\right)$

(20)

Now we see explicitly the fact that the turnaround times for the quark and antiquark, which were the same in the lab frame, are different in the moving frame because these events take place at different points in space. The simplest way to understand how this actually enables the object to move is to see it in the animation, STRINGY.

We can also understand what is happening in (20) by thinking about the forces involved. At $t = 0$ in the moving frame, while there is no apparent motion of the system because the quark and antiquark are moving at the speed of light in opposite directions, they do have different momenta. The quark momentum is red-shifted, because it is moving away from observers in the moving frame, to

$$\kappa\sqrt{\frac{1-v}{1+v}} = (1-v)\,\gamma\,\kappa \tag{21}$$

The antiquark momentum is blue-shifted, because it is moving towards observers in the moving frame, to

$$\kappa\sqrt{\frac{1+v}{1-v}} = (1+v)\,\gamma\,\kappa \tag{22}$$

The forces on these two quarks in the moving frame are **exactly the same as in the lab frame!** The reason, as we discussed earlier, is that the relativistic string is completely featureless. Looking at a straight piece of relativistic string, if you cannot see the ends, you cannot tell whether it is moving or not. Thus the force on the particles is $\pm T$. If you don't believe this argument, you can verify that it is correct (much more laboriously) by understanding how to Lorentz transform forces. The result is that the component in the direction of the Lorentz transformation is unchanged, while the components transverse to the direction are reduced by a factor of $1/\gamma$ (we will see this more explicitly later). The consequence of all this is that the quark in the moving frame loses all its energy and turns around in a time $(1-v)\gamma\kappa/T$, in agreement with (20), while the antiquark, which began with more energy, doesn't lose it all and turn around until $(1+v)\gamma\kappa/T$. In the intervening time, while both quark and antiquark are moving the $-x$ direction, the string continues to pull in the $+x$ direction on the antiquark and in the $-x$ direction on the quark, so energy is continuously transferred from the antiquark to the quark until the antiquark turns around. I like to call this system a relativistic push-me-pull-you, for obvious reasons.

Note that the comment about the "center of mass" on page 8 was designed only to confuse you. Center of mass has no significance in a relativistic system. What matters is center of energy. You will explore this further on the problem set.

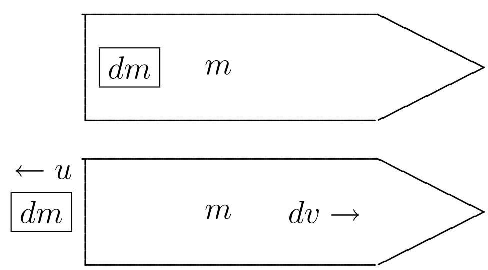
lecture 15

Topics:

- Rocket motion
- Review of circular motion
- Spinning relativistic string
- Relativistic traffic

Rocket motion

The surprising thing (to me) about relativistic rocket motion is that there is a sense in which it looks just the same as nonrelativistic rocket motion - at least in the rest frame of the rocket. Nonrelativistic rocket motion looks like this if the fuel is ejected at speed u:



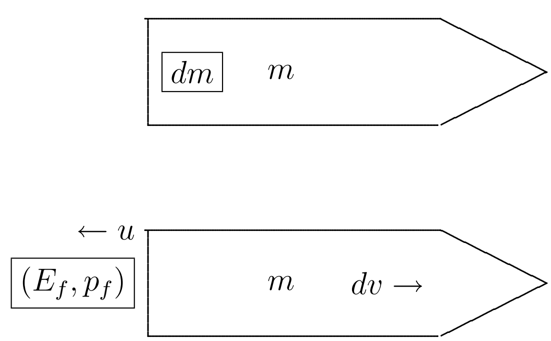
$$p_r = 0 \quad t = 0 \quad p_f = 0$$

$$p_r = m\,dv \quad t = dt \quad p_f = -dm\,u$$

momentum conservation ⇒

$$\frac{dp}{dt} = m\frac{dv}{dt} = \frac{dm}{dt}u = \text{Force} \tag{1}$$

Relativistic rocket motion looks like this:



$$E_r = m \quad p_r = 0 \quad t = 0 \quad p_f = 0 \quad E_f = dm$$

$$E_r = m \quad p_r = m\,dv \quad t = dt \quad p_f = u\,dm \quad E_f = dm \quad \text{NOT } dm\gamma_u$$

momentum conservation ⇒

$$\frac{dp}{dt} = m\frac{dv}{dt} = \frac{dm}{dt}u = \text{Force} \tag{2}$$

There are two compensating differences. Mass is no longer conserved in the relativistic situation, so it is not correct to say that the decrease in mass of the rocket is equal to the mass of the fuel ejected. But it **is** correct to use energy conservation and say that the decrease in energy of the rocket (which in the rest frame is just the decrease in mass) is equal to the energy of the fuel ejected. But then the momentum of the fuel ejected is not u times the mass, but rather u times the energy, which is the same as we got in the nonrelativistic case.

There are lots of differences that emerge as the velocity of the rocket becomes large, but these are entirely the result of time dilation and the different relation in relativity between the momentum p of the rocket and its speed, v.

Review of circular motion

You have all learned in previous physics courses about uniform circular motion. I thought that it would be useful to review this before going on to discuss rotations in more generality. So let us consider an object that is moving in a circle of radius R centered at the origin with a constant speed, v. This is motion in a single plane (because every circle lies in some plane), so it is convenient to choose our coordinate system so that the motion is in the x-y plane with $z = 0$. We can also choose to have the motion in the counterclockwise direction as seen from above (by choosing which is the positive z direction). Then the motion of our particle can be written as

$$x(t) = R\cos(\omega t + \phi) \qquad y(t) = R\sin(\omega t + \phi) \qquad z(t) = 0 \tag{3}$$

where ω is the angular velocity in radians per unit time. This defines the position vector of the particle, $\vec{r}\,(t)$, in the usual way:

$$\vec{r}\,(t) = x(t)\,\hat{x} + y(t)\,\hat{y} + z(t)\,\hat{z} = \hat{x}\,R\cos(\omega t + \phi) + \hat{y}\,R\sin(\omega t + \phi) \tag{4}$$

This should be familiar. We saw it earlier in the course in another context when we were discussing the connection between uniform circular motion and complex exponential. This motion in the x-y plane is precisely the same as the motion of the complex exponential

$$R\,e^{i\omega t} \tag{5}$$

in the complex plane.

The velocity of the particle is the derivative

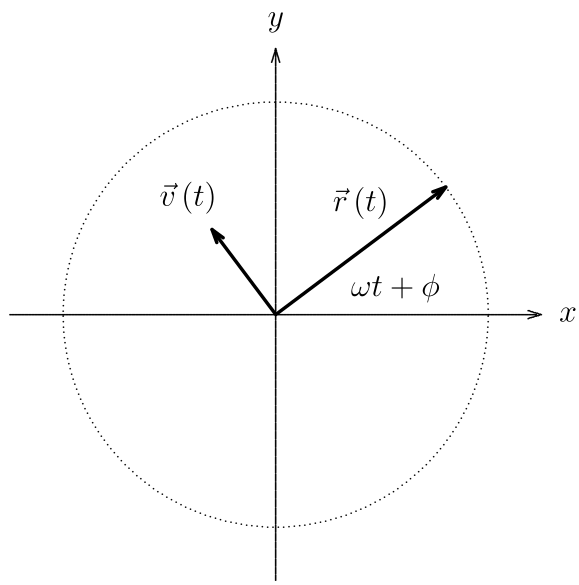
$$\vec{v}\,(t) = \dot{\vec{r}}\,(t) = -\hat{x}\,R\omega\sin(\omega t + \phi) + \hat{y}\,R\omega\cos(\omega t + \phi) \tag{6}$$

Note that because $\sin^2\theta + \cos^2\theta = 1$, the length of $\vec{v}\,(t)$ is $R\omega$, so sure enough, the speed of the particle is constant,

$$v = |\vec{v}\,(t)| = R\omega\,, \tag{7}$$

even though the velocity vector is constantly changing.

The relation between (3) and (6) is illustrated below at time t (for $\omega > 0$):



(8)

Note that $\vec{v}\,(t)$ is always perpendicular to $\vec{r}\,(t)$ and still in the x-y plane (you can see this from the diagram, but you can also see it by explicitly computing the dot product, $\vec{r}\,(t) \cdot \vec{v}\,(t)$ and showing that it is zero for all t).[1]

We can compute the acceleration of the particle by differentiating (6).

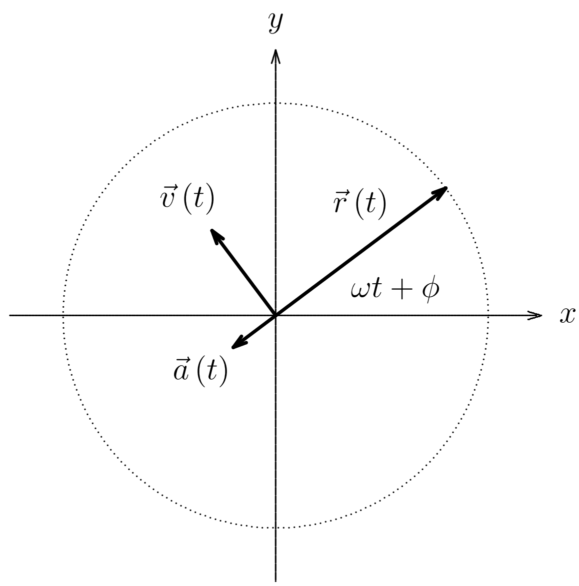
$$\vec{a}\,(t) = \dot{\vec{v}}\,(t) = -\hat{x}\,R\omega^2\cos(\omega t + \phi) - \hat{y}\,R\omega^2\sin(\omega t + \phi) = -\omega^2\;\vec{r}\,(t)\,. \tag{9}$$

The accleration in (9) is always directed to the center of the circle and the magnitude of the acceleration is $\omega^2 R$, or, using (7), this gives the (I hope familiar) formula

$$a = |\vec{a}\,(t)| = \omega^2 R = \frac{v^2}{R} \tag{10}$$

[1]In fact, $\vec{r}\,(t) \cdot \vec{v}\,(t) = 0$ for any motion on a circle, centered at the origin, whether it is uniform or not. You can see this by differentiating $\vec{r}\,(t) \cdot \vec{r}\,(t) = R^2$.

Here is (8) again with the acceleration shown this time.



(11)

The moral of this is much more general than it appears. Note that for uniform circular motion, as we go from $\vec{r}$ to $\vec{v} = \dot{\vec{r}}$ to $\vec{a} = \dot{\vec{v}} = \ddot{\vec{r}}$, each time we turn the vector counterclockwise by 90° and multiply by ω. We have thought about this in the context of a position vector. But any other vector (as we have already seen with $\vec{v}$) that undergoes uniform circular motion behaves the same way. We can get the time derivative by turning counterclockwise by 90° and multiplying by ω. This is telling us something very deep and important not just uniform circular motion, but about rotations and motion in general. We will begin to explore it in more generality in the next lecture.

Here is a simple application that is related to something that we know in various other ways. Suppose that we have a force of the form

$$\vec{F}(r) = -\alpha\, r^\beta\, \hat{r} \tag{12}$$

directed toward the origin with magnitude r^β. For what value of β do the circular orbits of a mass subject to this force have the same angular frequency independent of r? We want

$$\vec{F}(r) = -\alpha\, r^\beta\, \hat{r} = m\vec{a} = -m\,\omega^2\, r\, \hat{r} \tag{13}$$

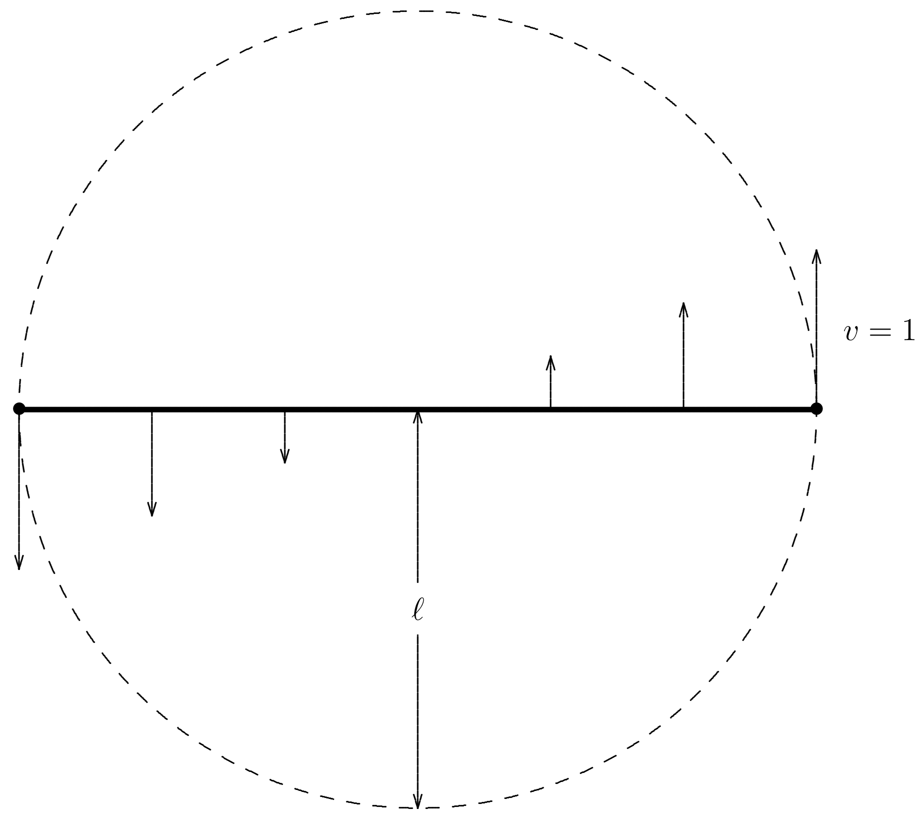
so

$$\omega^2 = \frac{\alpha}{m}\, r^{\beta-1} \tag{14}$$

and if $\beta = 1$ the angular frequency is independent of r. This is just a harmonic oscillator in more than one dimension.

Spinning relativistic string

Next time, we are going to start talking more seriously about the concept of angular momentum. We will find that it arises from the general principle of rotation invariance, and as such applies just as well to relativistic systems as in nonrelativistic mechanics. Indeed, one of the ways we know that relativistic string exists is by studying systems of quarks and antiquark that carry angular momentum. Let us study such a system. Consider a system of a massless quark and antiquark connected by relativistic string with tension T in its rest frame. Let us determine the conditions under which this system can rotate with the string straight and the quark and antiquark moving in a circle of radius ℓ as illustrated below:



(15)

The first thing to note is that the quark and antiquark must move at the speed of light because they are massless, by assumption. This has three immediate consequences. One is that the angular velocity of the system is (n relativistic units)

$$\omega = \frac{1}{\ell} \tag{16}$$

The second is that the quark and antiquark can carry only negligible momentum. This is because the force from the string at the end is Lorentz transformed to zero. You saw in Morin's book that the force from a string moving in the transverse direction with velocity v is reduced to $T\sqrt{1-v^2}$. This effect is simply time dilation. The transverse components of momentum do not change under a Lorentz transformation. Because we see moving clocks tick slowly by a factor of $1/\gamma$, we also see the rate of change of transverse momentum reduced by a factor of $1/\gamma$. In this case, where the quark and antiquark at the end are massless and move at the speed of light, $1/\gamma = 0$ and thus the rate of change of the momentum of the quark and antiquark must vanish. But the rate of change of momentum of a particle with momentum of magnitude p in uniform circular motion is $p\omega$. Because $\omega \neq 0$, we must have $p = 0$. Thus all the momentum and angular momentum in this system is carried by the string.

The third consequence is that at a point on the string a distance r from the center is moving with velocity r/ℓ, as shown in (15).

Next, let us check that the system can hold itself together. The force from the string at some radius r_0 must be right to produce the appropriate rate of change of momentum, which is the change of momentum associated with the uniform circular motion of the part of the string for $r > r_0$. To see that this works, let's calculate the momentum carried by the portion of the string for $r > r_0$. The contribution from an infinitesimal bit of string between r and $r + dr$ is

$$\frac{dm}{\sqrt{1-v^2}}\, v = \frac{T\, dr}{\sqrt{1 - r^2/\ell^2}}\, \frac{r}{\ell} \tag{17}$$

Then the momentum carried by the string for $r > r_0$ is

$$\int_{r_0}^{\ell} \frac{T\, dr}{\sqrt{1 - r^2/\ell^2}}\, \frac{r}{\ell} \tag{18}$$

If we change variables to $z = r/\ell$, this becomes

$$T\ell \int_{r_0/\ell}^{1} \frac{z\, dz}{\sqrt{1-z^2}} = -T\ell \int_{r_0/\ell}^{1} dz\, \frac{d}{dz}\sqrt{1-z^2} = T\ell\sqrt{1 - r^2/\ell^2} \tag{19}$$

The rate of change of momentum is then

$$p\omega = p/\ell = T\sqrt{1 - r_0^2/\ell^2} \tag{20}$$

This is right, because it reproduces the string force at r_0

Now that we know how this object moves and that it hangs together, we can ask what its energy and angular momentum are (it is at rest, so the total 3-momentum is zero). The energy from a small bit of string is just like (17) but without the factor of r/ℓ at the end. The total energy is then

$$\int_{-\ell}^{\ell} dr\, \frac{T\, dr}{\sqrt{1 - r^2/\ell^2}} = T\ell \int_{-1}^{1} \frac{dz}{\sqrt{1-z^2}} \tag{21}$$

This should be a familiar integral — we can make the trigonometric substitution

$$z = \sin\theta \qquad\qquad dz = d\theta\cos\theta \tag{22}$$

to write it as

$$E = T\ell \int_{-\pi/2}^{\pi/2} d\theta = \pi T\ell \tag{23}$$

To compute the angular momentum, we need an additional factor of $rv = r^2/\ell$ in the integrand of (21):

$$\int_{-\ell}^{\ell} dr\, \frac{T\, r^2\, dr}{\ell\sqrt{1 - r^2/\ell^2}} = T\ell^2 \int_{-1}^{1} \frac{z^2\, dz}{\sqrt{1-z^2}} \tag{24}$$

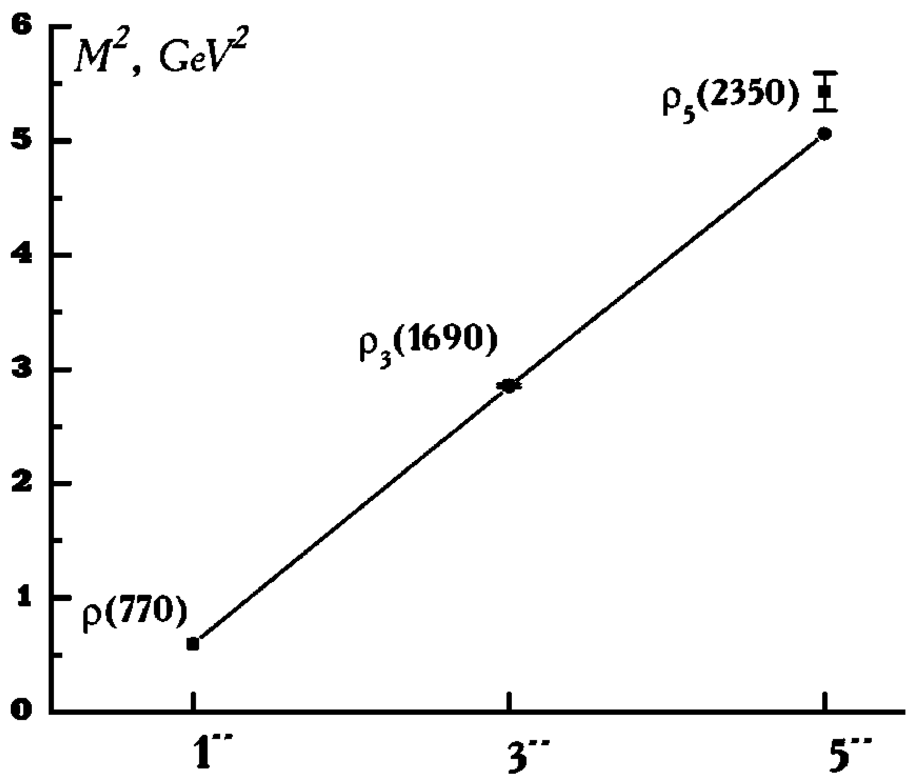
The same trigonometric substitution allows us to write this as

$$L = T\ell^2 \int_{-\pi/2}^{\pi/2} d\theta \, \sin^2\theta = \frac{1}{2}\pi T\ell^2 \tag{25}$$

Now the point is that we can eliminate ℓ from (23) and (25) and write this as a relation between the energy and angular momentum

$$E^2 = \pi^2 T^2 \ell^2 = 2\pi T \, L \tag{26}$$

Sure enough, when we look at quark-antiquark bound states, we see families of particles with similar properties in which E^2 grows linearly with L within each family.[2] Here is some representative data:



The data shows a slope of about

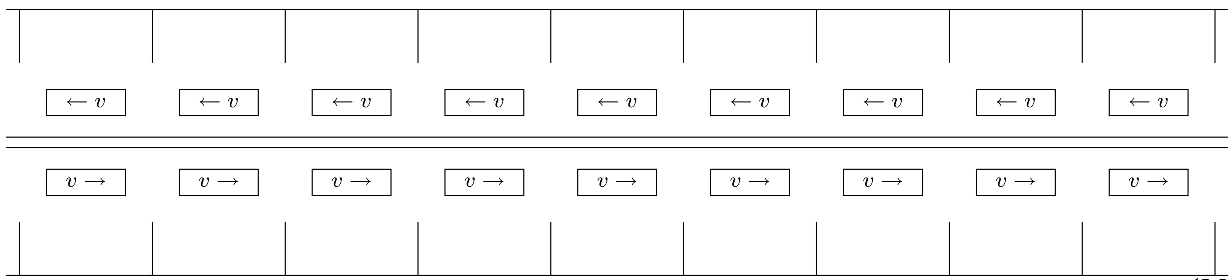
$$2\pi T \approx \frac{1 \text{ GeV}^2}{\hbar} \tag{27}$$

From this we can extract an approximate value for T, which works out to about 15 tons! These little particles are pulled by a rather strong force! That's why quarks and antiquarks don't get very far apart!

[2] See for example http://arxiv.org/abs/hep-ph/0103274.

Relativistic traffic

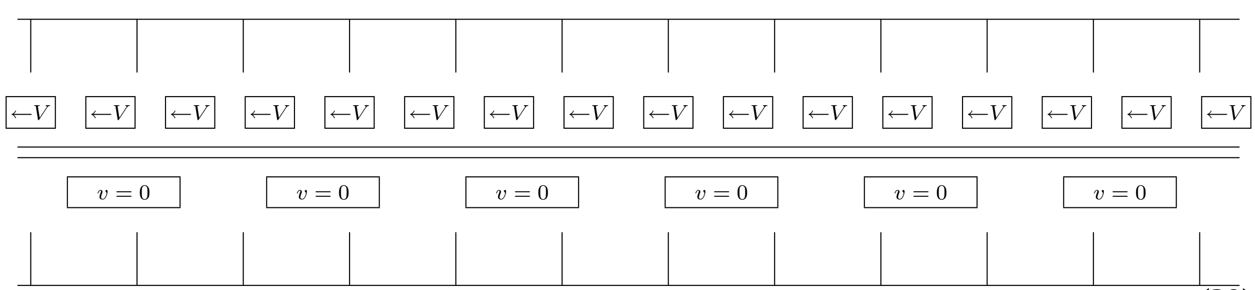
Consider relativistic traffic on a Boston street. We will make the completely unrealistic assumption that the traffic is flowing smoothly, with evenly spaced cars moving in opposite directions at speed v as shown below.



(28)

You will notice that there is something else that is unrealistic about this picture. There are plentiful parking spaces on the side of the road. This picture is animated in the *Mathematica* program **traffic.nb**.

Now let us look at this picture from the inertial frame of a driver in the lower lane of traffic. In the animation, you can change the inertial frame gradually. Try increasing the speed of the frame gradually, until you to get to an inertial frame that is moving along with the lower cars. Of course, as you increase the speed of the inertial frame, the parking spaces appear to move in the opposite direction, and the cars in the other lane move by faster. But here, we also see a number of effects of Lorentz contraction. The animation assumes that the initial speed of the cars is $0.9c$, so that things are very relativistic and Lorentz contraction is a big effect (and likewise the cars moving in the opposite direction don't move much faster because they were already moving at close to c in the original frame. I have done this to make the important effects show up better. In the actual physical demo you will see, the effect is much smaller, but the qualitative message is the same. Notice that as the lower cars slow down in the moving frame, they also appear longer. This is because they appeared Lorentz contracted in the initial frame, and going to the moving frame undoes some of the effect. The parking spaces and the upper cars going in the opposite direction are Lorentz contracted more and more as the speed of the frame increases. Finally, we get to the frame in which the lower cars are at rest, which looks sort of like this:

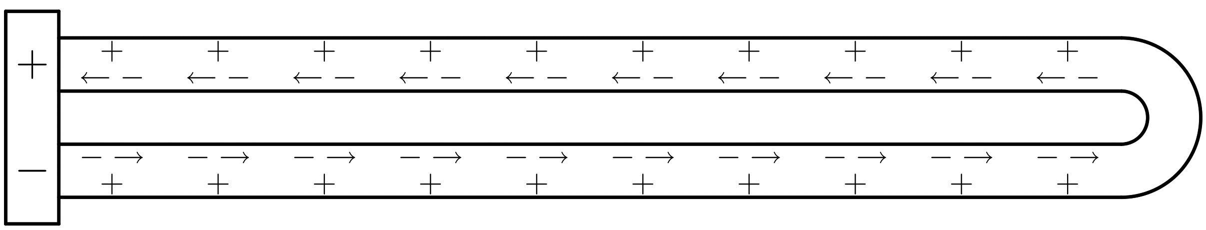


(29)

There are several things to notice about this. One is you see all the other cars in your lane farther apart than they were in the frame of the road, and they don't seem to fit into the parking spaces anymore. This is just Lorentz contraction. This is the same thing that is going on the problem of the relativistic chocolate chip cookies in David Morin's book. There you remember that the cookie dough was moving at relativistic speed on a conveyor belt and a circular cookie press was stamping out cookies - but the because the dough is Lorentz contracted, the cookies come out elongated. The cars going in the other direction are Lorentz contracted even more, and look shorter and closer together.

Something may be bothering you about this. How can the density of parking spaces change relative to the density of the cars? Where do those extra parking spaces come from? If we can harness this, we will be rich beyond the dreams of avarice. In fact, we can't make any money on it, because these extra parking spaces are coming in from the past and the future. This is possible everywhere if we have an infinite line of them stretching out in both directions, but it is always possible in some finite region of the roadway. The density in the rest frame is not the same as the density in a moving frame because of the relativity of simultaneity (which is the culprit in most such puzzles). But again, the effect is even larger for the cars going in the opposite direction. Because they are moving faster, they are even more Lorentz contracted than the parking spaces.

Of course, I don't really have any relativistic cars available to test these theories. But there is a very interesting analogy. Instead of two opposing lanes of traffic, consider current in a wire flowing up and back from a battery.



In this analogy, the negatively charged electrons are the cars and the rest of the electrical structure of the wire form positively charged parking spaces. Of course the actual motion is more chaotic — more like real Boston traffic. But the net current is as shown. The whole system looks electrically neutral because there is one car for each parking space. Now the question is, what forces do the components of one of the wires feel do to the moving electrons in the other. We expect from the lab frame where everything is neutral that there is no force on the fixed positive matrix - the parking spaces. And since there is no force in the lab frame, there is no force in any frame. But what force do the electrons feel. Now it is not so obvious, because the electrons are moving in the lab frame. So even though everything is neutral, there might still be a force that depends on velocity. Thus it is simplest to think about it in the inertial frame moving along with the electrons in one of the wires. In this frame, any force on the electrons proportional to velocity goes away, because they are at rest. But in this frame, as the animation or (29) shows, the electrons in the lower wire see far more negatively charge electrons in the upper wire than they see positively charged parking spaces for the electrons. Thus they experience a repulsive force. In the original frame, there is no Coulomb force because everything is neutral. In this frame, the repulsive force is still there, but it is interpreted as magnetism. When I connect the battery and get my two lanes of traffic moving, there should be a repulsive force between them.

Thus as you will see in more detail in Physics 15b, magnetism is a relativistic effect. When you feel the mysterious force between two magnets, that I am sure has fascinated all of you at one time or another, you are feeling in your bones the effect of special relativity.

lecture 16

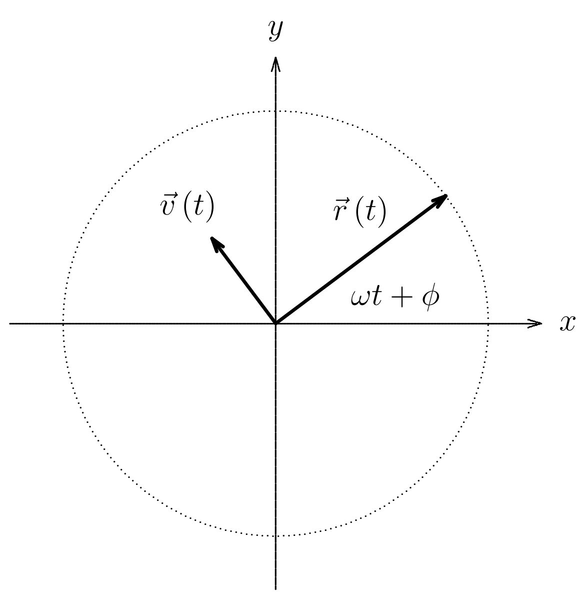
Topics:

Where are we now?

Now we are going to change gears and talk about angular momentum. We will see that conservation of angular momentum is associated with the invariance of the laws of physics under rotations. In that sense, the subject of angular momentum is not so different from relativity. Both are related to symmetries of the laws of physics.

Cross products - introduction

Last time, we reviewed uniform circular motion. To get a deeper understanding of circular motion in three dimensions, and to ready ourselves to study the more general problem of rigid body rotations, we will reformulate the discussion of uniform circular motion in terms of cross products. The advantage of this reformulation is that it makes the vector structure of the rotations more explicit. This is what we will need to take the next steps. Cross products are going to be particularly convenient as we talk about rotations and angular momentum.



(1)

Look at (1), from last time. As we go from the position, $\vec{r}$, to the velocity, $\vec{v}$, to the acceleration, $\vec{a}$, each time I just rotate the vector by $90°$ and change the length by a factor. We can rewrite this in a very useful way using cross products. The cross product can be defined by

$$\left[\vec{A}\times\vec{B}\right]_x = A_yB_z - A_zB_y\,, \quad \left[\vec{A}\times\vec{B}\right]_y = A_zB_x - A_xB_z\,, \quad \left[\vec{A}\times\vec{B}\right]_z = A_xB_y - A_yB_x\,. \quad (2)$$

This looks complicated (I hope that it will seem simpler after you have read the discussion in the appendix). The key point is that the formula for the velocity can be written as

$$\frac{d}{dt}\,\vec{r}\,(t) = \dot{\vec{r}}\,(t) = \omega\,\hat{z}\times\,\vec{r}\,(t) \quad (3)$$

because $\vec{r}\,(t)$ satisfies

$$\vec{r}\,(t)\cdot\hat{z} = 0\,. \quad (4)$$

The $\hat{z}\times$ produces the $90°$ twist, and the factor of ω rescales the vector by a factor (necessary in this case just to get the dimensions right).

You can see from the explicit definition in the appendix the following property:

> If $\vec{\alpha}$ is in the x-y plane, then $\hat{z}\times\vec{\alpha}$ is also in the plane, it has the same magnitude as $\vec{\alpha}$, and its direction is that of $\vec{\alpha}$ rotated by $90°$ counterclockwise about the z axis. (5)

This is just the right-hand rule. If $\vec{\alpha}$ is not in the x-y plane (as you can see in more detail in the appendix) then the cross product just throws away the component that is not in the plane and acts in the usual way according to (5) of the component in the place.

The relation (3) is very general. Not only is the time derivative given by this cross product for the position vector, but any vector undergoing uniform circular motion in the x-y plane behaves the same way.

> For any vector $\vec{u}\,(t)$ undergoing uniform circular motion in the x-y plane, the time derivative is given by the cross product (6)
>
> $$\frac{d}{dt}\vec{u}\,(t) = \omega\hat{z}\times\vec{u}\,(t)\,.$$

In particular, this gives us another way of calculating the acceleration. The velocity vector is also undergoing uniform circular motion. Thus

$$\vec{a}\,(t) = \omega\,\hat{z}\times\vec{v}\,(t) = \omega\,\hat{z}\times\left(\omega\,\hat{z}\times\,\vec{r}\,(t)\right) = -\omega^2\,\vec{r}\,(t) \quad (7)$$

Where the last step follows because each cross product with $\hat{z}$ just rotates $\vec{r}$ by $90°$, so the two cross products give a rotation by $180°$, which is just a sign change.

One of the crucial things about describing uniform circular motion in this way with cross products is that (3), (4) and (7) are vector equations. The $\hat{z}$ appears in (3), (4) and (7) because rotation in the x-y plane is rotation about the z axis. But because these are vector equations, we can rotate them to an arbitrary coordinate system. Then we can describe uniform circular motion in any plane through the origin, about an arbitrary axis through the origin. If the rotation is in some other plane through the origin, say perpendicular to a unit vector $\hat{n}$ (it is then rotation about the axis, $\hat{n}$), then the rotated versions of (3), (4) and (7) are

$$\vec{v}\,(t) = \frac{d}{dt}\,\vec{r}\,(t) = \omega\,\hat{n} \times \vec{r}\,(t)\,, \tag{8}$$

$$\vec{r}\,(t) \cdot \hat{n} = 0 \tag{9}$$

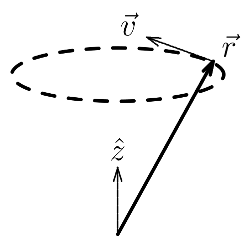
and

$$\vec{a}\,(t) = \frac{d}{dt}\,\vec{v}\,(t) = \omega\,\hat{n} \times \Big(\omega\,\hat{n} \times \vec{r}\,(t)\Big) = -\omega^2\,\vec{r}\,(t)\,, \tag{10}$$

where the last step follows because the cross product with $\hat{n}$ on a vector perpendicular to it just produces a rotation of 90° in the plane perpendicular to $\hat{n}$. Remember that (8) works only for motion that is circular! But still as we will see this is very powerful.

More general motion about an axis

Let's go back to circular motion about the z axis and think briefly about what happens if the vector r is not in the x-y plane. In this case, uniform circular motion about the z axis looks something like this:



(11)

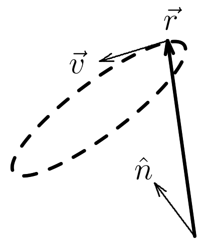
The velocity is in the x-y plane. What is going on is that the component of $\vec{r}$ perpendicular to $\hat{z}$ is undergoing uniform circular motion in the x-y plane while nothing is happening to the component of $\vec{r}$ in the $\hat{z}$ direction. This is very elegantly encoded in the cross product formula, (3) if we just drop the condition (4) that $\vec{r}$ be perpendicular to $\hat{z}$ and thus in the x-y plane. Thus

$$\vec{v}\,(t) = \omega\,\hat{z} \times \vec{r}\,(t) \tag{12}$$

is more general than we claimed above. Furthermore, there is really no reason to restrict ourselves to uniform circular motion. If at some time t, the vector $\vec{r}$ is instantaneously rotating about the z axis with angular velocity $\omega(t)$, the velocity is

$$\frac{d\vec{r}}{dt} = \vec{v}\,(t) = \omega(t)\,\hat{z} \times \vec{r}\,(t) \tag{13}$$

Finally, again because (13) is a vector equation, we can describe instantaneous rotation about an arbitrary axis $\hat{n}$ by just replacing $\hat{z} \to \hat{n}$:



(14)

Thus we can make the following very general and very important statement:

If at some time t, the vector $\vec{r}$ is instantaneously rotating about the $\hat{n}$ axis through the origin with angular velocity $\omega(t)$, the velocity is (15)

$$\frac{d\vec{r}}{dt} = \vec{v}\,(t) = \omega(t)\,\hat{n} \times \vec{r}\,(t)$$

This is a very general connection between a rotation and a cross product. Eventually, we will discuss what happens if the rotation is about an axis that does not go through the origin - but this is enough for now. We will use (15) very often so make sure that you understand it.

Small rotations

It will also be useful to rewrite (15) as an expression for the instantaneous change in a vector under an infinitesimal rotation, which we can do by noting that

$$\omega(t) = \frac{d\theta}{dt} \tag{16}$$

so that multiplying both sides of (15) by dt gives $d\vec{r}$, the infinitesimal change in the vector $\vec{r}$ under an infinitesimal rotation $d\theta$,

$$d\vec{r} = d\theta\,\hat{n} \times \vec{r} \tag{17}$$

One reason for thinking about infinitesimal rotations is that we know that infinitesimal symmetry transformations lead to conserved quantities by Noether's theorem. For infinitesimal rotations, the corresponding conserved quantity is a component of the angular momentum.

For a 3-dimensional system described by some number of 3-vector coordinates $\vec{r}_j$, the conserved quantity associated with invariance under rotations about an axis $\hat{n}$ is

$$\vec{L} \cdot \hat{n} \tag{18}$$

where

$$\vec{L} = \sum_j \vec{r}_j \times \vec{p}_j \tag{19}$$

where $\vec{p}_j$ is the momentum of the jth particle,

$$\vec{p}_j = \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_j} \tag{20}$$

If the Lagrangian is invariant under all rotations, all components, and therefore the entire vector, $\vec{L}$ is conserved — $\vec{L}$ is the angular momentum — this is angular momentum conservation.

What I think is worth understanding about this is the way that the appearance of the cross product in rotations leads to the appearance of the cross product in the expression for the angular momentum. This may have seemed mysterious to you, even as you followed the derivations in previous physics courses showing that angular momentum is conserved under certain situations. What is really going on is conservation of angular momentum appears naturally from Noether's theorem in a Lagrangion that is invariant under rotations. The mysterious cross product comes simply from the mathematical description of infinitesimal rotations in 3-dimensional space.

The reference point

So far we have found the velocity of a vector rotating around an axis that goes through the origin of the coordinate system. We used this to find the change in a vector due to a small rotation around an axis through the origin and from this we derived the form of the conserved angular momentum that follows from invariance under such a rotation. I now want to remove the restriction that the axis goes through the origin, so that we know what happens for a totally arbitrary rotation. This is not difficult. If the axis goes through some arbitrary point $\vec{R}$ and the axis points in the direction $\hat{n}$, the velocity of the point $\vec{r}_j$ is given by

$$\omega\, \hat{n} \times (\vec{r}_j - \vec{R}\,) \tag{21}$$

This must be right, because subtracting $\vec{R}$ from every vector in the original coordinate system does not affect the velocities (because $\vec{R}$ is fixed) and takes us to a coordinate system in which $\vec{R}$ is the origin, in which (21) is equivalent the result we derived above. Notice that because $\hat{n} \times \hat{n} = 0$, we can add to $\vec{R}$ any multiple of $\hat{n}$ without affecting (21):

$$\omega\, \hat{n} \times (\vec{r}_j - \vec{R}\,) = \omega\, \hat{n} \times (\vec{r}_j - \vec{R} - \alpha\hat{n}\,) \tag{22}$$

Following the same argument that we gave above, we see that if a Lagrangian $\mathcal{L}(\vec{r}, \dot{\vec{r}}\,)$ is invariant under rotations about an axis through the point $\vec{R}$ in the $\hat{n}$ direction, then Noether's theorem implies that there is a conserved quantity of the form

$$\hat{n} \cdot \vec{L}_{\vec{R}} = \hat{n} \cdot \Big((\vec{r} - \vec{R}\,) \times \vec{p} \Big) \tag{23}$$

and if $\mathcal{L}$ depends on several vectors, the conserved quantity is

$$\hat{n} \cdot \vec{L}_{\vec{R}} = \hat{n} \cdot \sum_j (\vec{r}_j - \vec{R}\,) \times \vec{p}_j \tag{24}$$

If the Lagrangian is invariant not just under rotations about the axis $\hat{n}$ through $\vec{R}$, but under rotation about **any** axis through $\vec{R}$, then the entire angular momentum vector is conserved,

$$\vec{L}_{\vec{R}} = (\vec{r} - \vec{R}\,) \times \vec{p} \tag{25}$$

and if $\mathcal{L}$ depends on several vectors, the conserved quantity is

$$\vec{L}_{\vec{R}} = \sum_j (\vec{r}_j - \vec{R}\,) \times \vec{p}_j \tag{26}$$

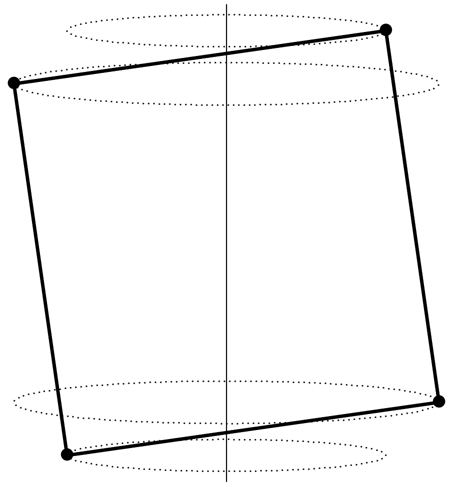
In this case, the point $\vec{R}$ is called the "reference point" and we talk about $\vec{L}_{\vec{R}}$ as "the angular momentum about the point $\vec{R}$." In this case, there is no ambigiuty in $\vec{R}$ (that is there is no analog of (22)). In (26) (and more trivially in (25)), you see that the $\vec{L}_{\vec{R}}$ is independent of $\vec{R}$ if the total momentum vanishes.

Rigid bodies rotating about a fixed axis

Now that we know from our study of relativity that rigid bodies don't exist, we are going to study them in detail. Of course, this is not quite as stupid as it sounds. What we mean by a hypothetical rigid body is one in which the shape is completely fixed. Mathematically, we can describe this by saying all the lengths of vectors between different parts of the system are fixed by the internal dynamics of the system. We know that this can only be an approximation. Whatever the internal dynamics is that keeps the system rigid, it will always be possible to deform the system slightly. But it can be a very good approximation if we are concerned only with relatively slow motions of the whole system. What we really mean when we say that a body is rigid is that all the modes that correspond to deformations of the system have very large frequencies, and/or very large damping, so that we can ignore them.

The most general continuous motions that preserve the shape of a rigid body are combinations of translations and rotations. This crucial fact is proved in David Morin's book, but I hope that it is obvious to you. In fact, we can always describe the position of a rigid body by giving the position of any convenient point on the body, and specifying the rotation required to get to its orientation from some standard orientation. For bodies that rotate only about a single axis, the situation is simpler, because we can specify the rotation by just giving the angle θ of the system from some fixed position. Thus a rigid body rotating about a fixed axis is a system of one degree of freedom.

Suppose that the fixed axis is in the $\hat{n}$ direction, and that the axis goes through the origin of our coordinate system. Then when the body rotates, every part of the rigid body executes circular motion about some point on the axis, and every part moves with the same angular velocity. The radius of the circular motion is the distance from the axis. This is illustrated below for a square made of balls and sticks:



(27)

The kinetic energy of such a rigid body is given by the sum of $\frac{1}{2}mv^2$ for each of the parts of the system.[1] We are going to compute the kinetic energy in terms of the angular velocity, $\dot{\theta}$, and discuss it. I will do this for an arbitrary axis $\hat{n}$, because I want you to get used to dealing with arbitrary axes. But because I also want you to understand what I am talking about, I will repeat some of the steps for the special case of rotations about the z axis — that is $\hat{n} = \hat{z}$.

Suppose that the system is made up of point masses with positions at some time t given by the vectors $\vec{r}_j = (x_j, y_j, z_j)$, and that the angular velocity of the system about $\hat{n}$ at time t is $\dot{\theta}$. Then the velocity of the point with position $\vec{r}_j$ is

$$\dot{\theta}\,\hat{n} \times \vec{r}_j \tag{28}$$

It is perpendicular to $\hat{n}$ and $\vec{r}_j$ because of the form of the cross product. For $\hat{n} = \hat{z}$, this looks like

$$\dot{\theta}\,\hat{z} \times \vec{r}_j = \dot{\theta}\,(-y_j, x_j, 0) \tag{29}$$

— the velocity is in the x-y plane, perpendicular to $\hat{z}$.

Thus the kinetic energy for rotations about the $\hat{n}$ axis has the form

$$\frac{1}{2}\sum_j m_j\,\dot{\theta}^2\,|\hat{n} \times \vec{r}_j|^2 = \frac{1}{2}\,I\,\dot{\theta}^2 \tag{30}$$

where

$$I \equiv \sum_j m_j\,|\hat{n} \times \vec{r}_j|^2 \tag{31}$$

For $\hat{n} = \hat{z}$, this looks like

$$I \equiv \sum_j m_j\,|\hat{z} \times \vec{r}_j|^2 = \sum_j m_j\,(x_j^2 + y_j^2) \tag{32}$$

The quantity I in (30) and (31) is called the "moment of inertia" about the axis $\hat{n}$. As you see, it summarizes the behavior of the body under rotations about the axis $\hat{n}$. The length

$$|\hat{n} \times \vec{r}_j| \tag{33}$$

[1] If the system is continuous, we must replace the sum by an integral, but this doesn't change anything important.

(as you can see immediately for $\hat{n} = \hat{z}$) is just the perpendicular distance of the jth point mass from the axis $\hat{n}$. This remains constant as the body rotates, so we do not have to specify the time dependence of this, and I is independent of time. It is a constant that just depends on the axis and the masses in the rigid body and their positions.

The Lagrangian describing the motion of this system will involve the kinetic energy, (30). If we differentiate this with respect to $\dot{\theta}$, we get the generalized momentum associated with the angle θ that specifies the orientation of the rigid body. This is

$$L = I\,\dot{\theta} \tag{34}$$

For a rigid body rotating about the $\hat{n}$ axis, the "momentum" L in (34) is just the component of the angular momentum about the axis $\hat{n}$. Let's prove this. In general, the angular momentum is

$$\vec{L} = \sum_j \vec{r}_j \times \vec{p}_j = \sum_j m_j\,\vec{r}_j \times \dot{\vec{r}}_j = \sum_j m_j\,\vec{r}_j \times \dot{\theta}\,(\hat{n} \times \vec{r}_j) \tag{35}$$

so the component in the $\hat{n}$ direction is

$$\begin{aligned} \hat{n} \cdot \vec{L} &= \hat{n} \cdot \sum_j m_j\,\vec{r}_j \times \dot{\theta}\,(\hat{n} \times \vec{r}_j) = \dot{\theta} \sum_j m_j\,\hat{n} \cdot (\vec{r}_j \times (\hat{n} \times \vec{r}_j)) \\ &= \dot{\theta} \sum_j m_j\,(\hat{n} \times \vec{r}_j) \cdot (\hat{n} \times \vec{r}_j) = \dot{\theta} \sum_j m_j\,|\hat{n} \times \vec{r}_j|^2 = I\,\dot{\theta} \end{aligned} \tag{36}$$

At the end of (36), we have used the cyclic property of the triple product ($\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{c} \cdot (\vec{a} \times \vec{b})$ — an important fact that keeps coming up in the physics of rotations). The point is that if the axis is fixed, it is only the component of angular momentum about the axis that matters. Whatever is fixing the axis can provide torques that change the other components of angular momentum anyway, so they are not very interesting. But the component in the direction of the fixed axis will be conserved unless there is some physics that depends on the angle θ, or some other degree of freedom that interacts with the rigid body.

Just to say this once more in a different way, if the rigid body is free to rotate about the fixed axis $\hat{n}$, then the Lagrangian is just (30) and (34) is the quantity that is conserved by virtue of Noether's theorem, or equivalently the fact that the Lagrangian is independent of θ. It is clearly related to the angular momentum, but it is not the angular momentum — just one component of the angular momentum. We will see later that in general the other components are not conserved in this kind of motion because the physics that is fixing the axis is supplying a torque. These relations are very important. We will come back to them next week in a more general context.

The three dumbbells

A very simple system that shows some of the bewildering complexity of rotational motion is a system of two equal masses m attached to one another by a string and free to rotate without friction on a horizontal plane about their common center of mass. The configuration of this system for fixed R just depends on the angle with angular velocity $\omega = \dot{\theta}$.

If the distance between the objects is $2R$, so that both masses are a distance R from the center of mass, the angular momentum is perpendicular to the plane of the rotation and equal to

$$L = \hat{z} \cdot \vec{L} = \hat{z} \cdot \left(\vec{r} \times \vec{p}\right) = 2mR^2\omega \tag{37}$$

so in this case the moment of inertia is

$$I = 2mR^2 \tag{38}$$

Now suppose the distance between the masses is variable - $R(t)$ - where the t dependence is imposed by pulling in or letting out the string. It is probably obvious that angular momentum is conserved in this case, but let's be pedantic about it and think about the Lagrangian. It is just the kinetic energy

$$\mathcal{L} = mR(t)^2\dot{\theta}^2 + m\dot{R}(t)^2 = \frac{1}{2}I\,\dot{\theta}^2 + m\dot{R}(t)^2 \tag{39}$$

Sure enough, the Lagrangian does not depend on θ, so the momentum associated with θ is conserved - and this is just L

$$L = \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = 2mR(t)^2\dot{\theta} = 2mR(t)^2\omega \tag{40}$$

So if, for example, $R(t)$ is decreasing, ω must be increasing to keep the angular momentum constant.

This is very familiar. This is the way an ice-skater gets spinning very fast. But it is easy to get confused. When you pull on the rope, you are pulling perpendicular to the direction of motion of the masses. How is it that this speeds up the masses? In this simple case, it is easy to see how everything works. While the rope is being pulled in, there is a component of the motion of the mass in the direction of the centripetal force, and thus the centripetal force feeds energy into the kinetic energy of the masses, increasing their speed. But still this simple example is a taste of some of the bizarre behavior of angular momentum in three dimensions. We will see a lot more of it in the next few weeks.

A simple way to implement this system is to hold a weight in each hand and stand on a frictionless turntable.

Appendix 1 - cross products - details

The cross product is essentially just an antisymmetric combination of two vectors. This antisymmetric combination of two vectors in interesting because it defines a plane, and planes are intimately connected with rotations. The particularly convenient thing about this combination in three dimensional space is that it behaves like another vector. The cross product is the mathematical statement of the fact the antisymmetric combination of two vectors in three dimensional space defines a plane which in turn defines another vector. The geometrical definition of the cross product is a good way to see that it behaves like a vector under rotations, so we will start with that. Then I will indicate how we can show that this geometrical definition is equivalent to a definition given in terms of components.

The geometrical definition is this:

> Given two vectors, $\vec{A}$ and $\vec{B}$, the object $\vec{A} \times \vec{B}$ is a vector with magnitude $\left|\vec{A}\right| \left|\vec{B}\right| \sin\theta$ where θ is the angle between $\vec{A}$ and $\vec{B}$ defined as a positive angle between 0 and π. The direction of $\vec{A} \times \vec{B}$ is perpendicular to the plane formed by $\vec{A}$ and $\vec{B}$ with the sign determined by the right-hand rule. (41)

With this definition, it is easy to understand why $\vec{A} \times \vec{B}$ behaves like a vector under rotations. The magnitude doesn't change under a rotation because $\left|\vec{A}\right|$, $\left|\vec{B}\right|$ and $\sin\theta$ are all unchanged. And the direction rotates properly because it is tied to the directions of $\vec{A}$ and $\vec{B}$.

It is crucial that the cross product $\vec{A} \times \vec{B}$ is antisymmetric in the two vectors $\vec{A}$ and $\vec{B}$,

$$\vec{A} \times \vec{B} = -\vec{B} \times \vec{A} \tag{42}$$

In the geometrical definition, this follows from the application of the right hand rule. If you interchange $\vec{A}$ and $\vec{B}$, the cross product changes direction because the right hand rule goes from $\vec{B}$ to $\vec{A}$ rather than from $\vec{A}$ to $\vec{B}$. This antisymmetry ensures that either the two vectors $\vec{A}$ and $\vec{B}$ define a plane or the antisymmetric combination vanishes. Then the fact that in three dimensional space, there is a unique direction perpendicular to a given plane allows us to turn the antisymmetric combination into a vector.

The geometrical definition, (41), is equivalent to the following component definition,

$$\left[\vec{A}\times\vec{B}\right]_x = A_y B_z - A_z B_y\,, \quad \left[\vec{A}\times\vec{B}\right]_y = A_z B_x - A_x B_z\,, \quad \left[\vec{A}\times\vec{B}\right]_z = A_x B_y - A_y B_x\,, \tag{43}$$

where we are using a notation for vector components in which

$$\left[\vec{A}\right]_x = A_x\,, \quad \left[\vec{A}\right]_y = A_y\,, \quad \left[\vec{A}\right]_z = A_z\,. \tag{44}$$

If you have not seen cross products before in your math courses, you should look carefully at the demonstration of this equivalence below. We will be using cross products a lot for the next couple of months, so you might as well get used to them.

To prove that (41) and (43) are equivalent, we first show that

$$\left(\vec{A} \times \vec{B}\right) \cdot \vec{A} = \left(\vec{A} \times \vec{B}\right) \cdot \vec{B} = 0 \tag{45}$$

We can do this by explicit calculation. For example,

$$\left(\vec{A} \times \vec{B}\right) \cdot \vec{A} = (A_y B_z - A_z B_y)A_x + (A_z B_x - A_x B_z)A_y + (A_x B_y - A_y B_x)A_z = 0 \tag{46}$$

The calculation for $\vec{B}$ is similar (as it must be because of the antisymmetry of the cross product - (46) just says that the dot product of the cross product with the first vector in the cross product vanishes - and because of antisymmetry the same must be true for the second vector in the cross product). Thus $\vec{A} \times \vec{B}$ is perpendicular to both A and B and therefore perpendicular to the plane they form, just as in the geometrical definition.

You can see that the magnitude of the object given by (43) is right by explicitly calculating its square.

$$\left(\vec{A}\times\vec{B}\right)\cdot\left(\vec{A}\times\vec{B}\right) = (A_yB_z - A_zB_y)^2 + (A_zB_x - A_xB_z)^2 + (A_xB_y - A_yB_x)^2 \tag{47}$$

$$= A_x^2B_y^2 + A_x^2B_z^2 + A_y^2B_x^2 + A_y^2B_z^2 + A_z^2B_x^2 + A_z^2B_y^2 - 2A_xB_x\,A_yB_y - 2A_xB_x\,A_zB_z - 2A_yB_y\,A_zB_z \tag{48}$$

If we add and subtract $A_x^2B_x^2 + A_y^2B_y^2 + A_z^2B_z^2$ to this, the positive terms can be factored into

$$\left(A_x^2 + A_y^2 + A_z^2\right)\left(B_x^2 + B_y^2 + B_z^2\right) \tag{49}$$

and the negative terms into

$$-\left(A_xB_x + A_yB_y + A_zB_z\right)^2 \tag{50}$$

so we can write (47) as

$$\left(\vec{A}\times\vec{B}\right)\cdot\left(\vec{A}\times\vec{B}\right) = \left|\vec{A}\right|^2\left|\vec{B}\right|^2 - \left(\vec{A}\cdot\vec{B}\right)^2 = \left|\vec{A}\right|^2\left|\vec{B}\right|^2(1-\cos^2\theta) = \left|\vec{A}\right|^2\left|\vec{B}\right|^2\sin^2\theta \tag{51}$$

Finally, you can see the right-hand rule by calculating an example, like $\hat{x}\times\hat{y} = \hat{z}$. Thus we have checked that the component definition (43) is equivalent to the geometrical definition (41).

We can use the cross product very efficiently to describe uniform circular motion about the origin. First notice that the velocity

$$\vec{v}\,(t) = \dot{\vec{r}}\,(t) = -\hat{x}\,R\omega\sin(\omega t + \phi) + \hat{y}\,R\omega\cos(\omega t + \phi) \tag{52}$$

can be written as

$$\vec{v}\,(t) = \omega\,\hat{z}\times\,\vec{r}\,(t) \tag{53}$$

where $\vec{r}\,(t)$ satisfies

$$\vec{r}\,(t)\cdot\hat{z} = 0\,. \tag{54}$$

Let's see this explicitly:

$$\begin{aligned}
\left[\vec{v}\,(t)\right]_x &= \omega\,[\hat{z}\times\,\vec{r}\,(t)]_x = \omega\left([\hat{z}]_y[\,\vec{r}\,(t)]_z - [\hat{z}]_z[\,\vec{r}\,(t)]_y\right) = -\omega\,[\,\vec{r}\,(t)]_y = -\omega\,R\sin(\omega t + \phi)\\
\left[\vec{v}\,(t)\right]_y &= \omega\,[\hat{z}\times\,\vec{r}\,(t)]_y = \omega\left([\hat{z}]_z[\,\vec{r}\,(t)]_x - [\hat{z}]_x[\,\vec{r}\,(t)]_z\right) = \omega\,[\,\vec{r}\,(t)]_x = \omega\,R\cos(\omega t + \phi)\\
\left[\vec{v}\,(t)\right]_z &= \omega\,[\hat{z}\times\,\vec{r}\,(t)]_z = \omega\left([\hat{z}]_x[\,\vec{r}\,(t)]_y - [\hat{z}]_y[\,\vec{r}\,(t)]_x\right) = 0
\end{aligned} \tag{55}$$

(4) is the statement that $\vec{r}\,(t)$ is in the x-y plane and (3) describes the effect of the circular motion. As we saw, what the cross product of the unit vector $\hat{z}$ is doing on the vector $\vec{r}$ in the x-y plane is just rotating the vector counterclockwise by $90°$.

Another useful quantity is the so-called "triple product"

$$\left(\vec{A}\times\vec{B}\right)\cdot\vec{C} = A_xB_yC_z - A_yB_xC_z + A_yB_zC_x - A_zB_yC_x + A_zB_xC_y - A_xB_zC_y\,. \tag{56}$$

You can see explicitly that this has the property of **complete antisymmetry**. It changes sign if any two of the vectors are interchanged. This means also that the triple product is "cyclic" —

$$\left(\vec{A}\times\vec{B}\right)\cdot\vec{C} = \left(\vec{B}\times\vec{C}\right)\cdot\vec{A} = \left(\vec{C}\times\vec{A}\right)\cdot\vec{B} \tag{57}$$

This is a good thing to remember.

lecture 17

Topics:

Where are we now
Central forces
Energy and angular momentum
$V(\infty) = 0$
The parallel axis theorem
Example - Planar bodies moving in a plane
Torque - fixed reference point
Example — oscillations of a hanging rod
Torque - moving reference point
Example — Rolling ring

Where are we now

Rotations are important for two reasons. The highfalutin theoretical reason is that the laws of physics are apparently invariant under rotations. This leads to the important conservation law of angular momentum. The down-to-earth practical reason is that things rotate — and when rigid objects rotate, things get complicated and understanding the structure of rotational motion is crucial to understanding what happens.

In the last lecture, we began by reviewing uniform circular motion. We used the connection between uniform circular motion and the cross product to find the velocity of a vector rotating around an axis through some arbitary point $\vec{R}$.

> If at some time t, the vector $\vec{r}$ is instantaneously rotating about the $\hat{n}$ axis through some arbitary point $\vec{R}$ with angular velocity $\omega(t)$, the velocity is (1)
>
> $$\frac{d\vec{r}}{dt} = \vec{v}\,(t) = \omega(t)\,\hat{n} \times \left(\vec{r}\,(t) - \vec{R}\right)$$

We used this to motivate the definition of angular momentum from Noether's theorem, and to define the moment of inertia of an object rotating about a fixed axis.

Today, after a quick review of central forces and the effective potential, we are going to go on with our dicussion of moments of intertia about fixed axes, talk about the parallel axis theorem and discuss torque.

Central forces

One of the most important examples of gravity in action is the solar system, where the sun is much more massive than all the other bodies, so the center of mass is essentially at the position on the sun. Then to a good approximation, all the other planets orbit the sun. The gravitational forces of the planets on each other are very small compared to the force of the sun and can be ignored to first approximation. This means that we can put the sun at the origin and look at the orbit of one planet at a time. This in turn is an example of a more general system with a central potential, with Lagrangian

$$\mathcal{L}(\dot{\vec{r}}, \vec{r}) = \frac{1}{2} m \, \dot{\vec{r}}^{\,2} - V(r) \tag{2}$$

where

$$r = \left|\vec{r}\right| \tag{3}$$

This is invariant under rotations about any axis through the origin so there is a conserved angular momentum,

$$\vec{L} = \vec{r} \times \vec{p} = m \, \vec{r} \times \dot{\vec{r}} = m \, \vec{r} \times \vec{v} \tag{4}$$

While we know formally that the angular momentum is conserved because of Noether's theorem, it is instructive to see how the conservation arises in this particular case. Using the product rule, we can write

$$\frac{d}{dt} \vec{L} = m \frac{d}{dt} \vec{r} \times \dot{\vec{r}} = m \, \dot{\vec{r}} \times \dot{\vec{r}} + m \, \vec{r} \times \ddot{\vec{r}} \tag{5}$$

The first term vanishes trivially because $\vec{A} \times \vec{A} = 0$ for any vector $\vec{A}$ because of the antisymmetry of the cross product. The second term vanishes because

$$m \, \ddot{\vec{r}} = \vec{F} = -\frac{\partial}{\partial \vec{r}} V(r) = -\frac{\partial r}{\partial \vec{r}} V'(r) = -\frac{1}{r} \, \vec{r} \, V'(r) = -\hat{r} \, V'(r) \tag{6}$$

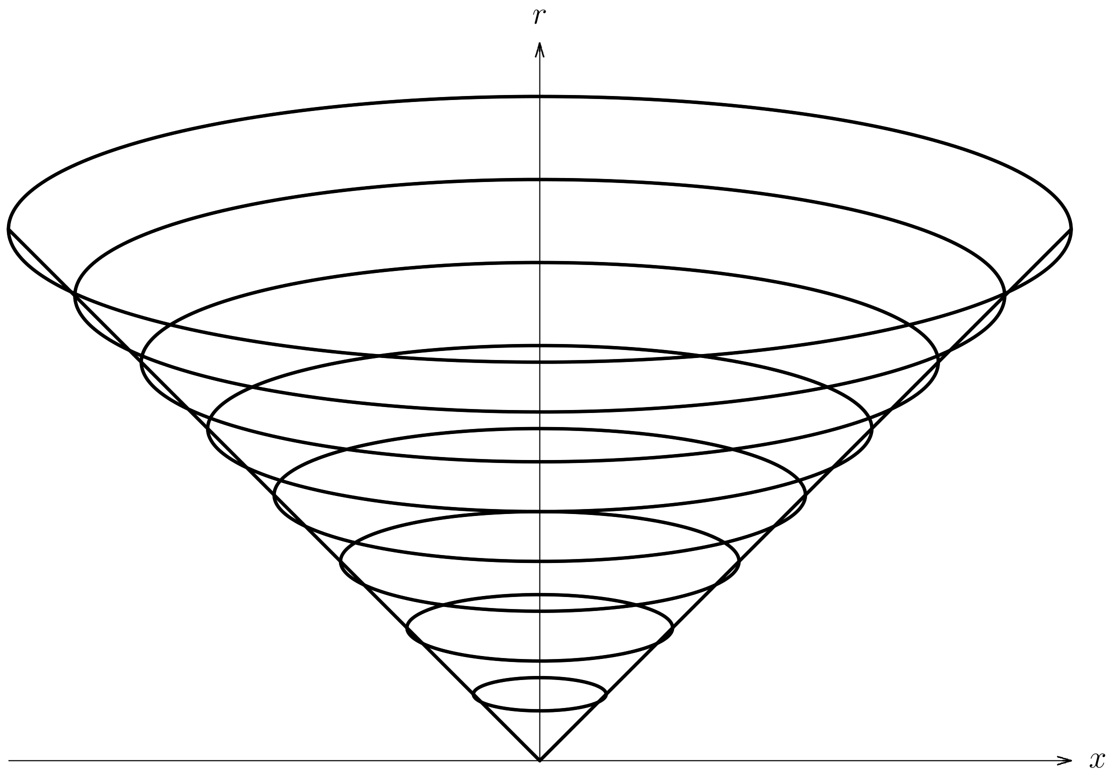
so the second term is proportional to $\vec{r} \times \vec{r} = 0$. The crucial step here is

$$\frac{\partial r}{\partial \vec{r}} = \frac{1}{r} \, \vec{r} \tag{7}$$

We have talked about this relation before, but it is so important that I want to give you several ways of remembering it. This can be seen in various ways. One of the nicest is to differentiate both sides of the equation $r^2 = \vec{r} \cdot \vec{r}$,

$$\frac{\partial}{\partial \vec{r}} r^2 = 2r \, \frac{\partial r}{\partial \vec{r}} = \frac{\partial}{\partial \vec{r}} \vec{r} \cdot \vec{r} = 2\vec{r} \tag{8}$$

from which (7) follows immediately. Another way of seeing this handy fact is to think about what r looks like considered as a function of the components of $\vec{r}$. This is easier to visualize in two dimensions, where we can plot $r = \sqrt{x^2 + y^2}$ versus x and y, and the result is a cone standing on its tip:



(9)

with the y axis going into the paper. Now $\frac{\partial}{\partial\vec{r}}$ is the gradiant operator $\vec{\nabla}$ that points in the direction in which the function it acts on is increasing, and has magnitude equal to the rate of increase. But the r function described by the cone increases linearly as one goes away from the origin in every direction, which is another way of seeing (7).

Energy and angular momentum

One of the most important consequences of angular momentum conservation for motion in a central force is that the motion is confined to a plane. This follows because $\vec{L} = \vec{r} \times \vec{p} = m\,\vec{r} \times \vec{v}$, it is perpendicular to both $\vec{r}$ and $\vec{v}$:

$$\vec{r} \cdot \vec{L} = \vec{v} \cdot \vec{L} = 0 \tag{10}$$

Therefore both the position vector and the velocity are always in the plane through the origin perpendicular to $\vec{L}$.

It is therefore very convenient to analyze the motion in polar coordinates in the plane of the motion. Let us rotate our coordinate system until the motion is in the x-y plane. The angular momentum is then in the z direction, and we can define our polar coordinates by

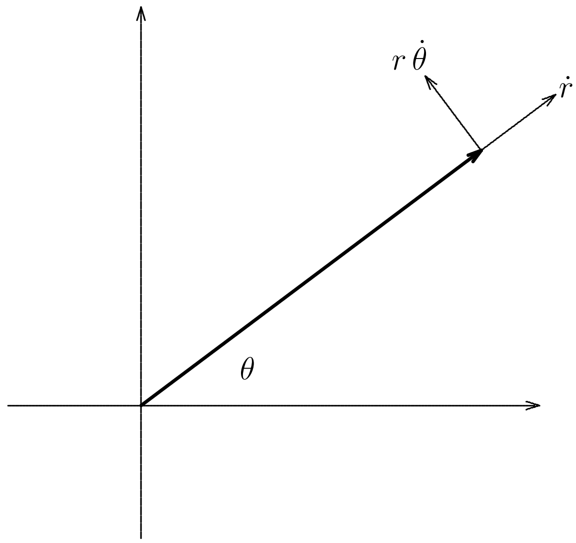
$$x = r\cos\theta \qquad y = r\sin\theta \tag{11}$$

In this coordinate system, the problem of finding the motion reduces to an equivalent problem in two dimensions. Because we know from the outset that z is going to be zero throughout the motion, we can ignore z completely and describe the system in terms of r and θ.

The square of the speed looks simple in both cartesian and in polar coordinates because

$$\dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2 \, \dot{\theta}^2 \tag{12}$$

which you can see either directly by differentiating (11) or by staring at the picture below and noting that the component of the velocity in the radial direction is perpendicular to the component of the velocity in the angular direction.



The Lagrangian for this equivalent two dimensional problem is

$$\mathcal{L}(\dot{r}, \dot{\theta}, r, \theta) = \frac{1}{2} m \, \dot{r}^{\,2} + \frac{1}{2} m \, r^{\,2} \, \dot{\theta}^{\,2} - V(r) \tag{13}$$

The Lagrange equation for $\ddot{r}$ is

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{r}} = \frac{\partial \mathcal{L}}{\partial r} \tag{14}$$

or

$$m \, \ddot{r} = m \, r \, \dot{\theta}^{\,2} - V'(r) \tag{15}$$

The Lagrange equation for $\ddot{\theta}$ is

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = \frac{\partial \mathcal{L}}{\partial \theta} \tag{16}$$

or

$$\frac{d}{dt} \, m \, r^2 \, \dot{\theta} = 0 \tag{17}$$

This expresses conservation of angular momentum in polar coordinates, because

$$m \, r^2 \, \dot{\theta} = m \, r \, v_\theta = [m \, \vec{r} \times \vec{v}]_z \tag{18}$$

Thus

$$m \, r^2 \, \dot{\theta} = L \qquad \qquad \dot{\theta} = \frac{L}{m \, r^2} \tag{19}$$

where L is the nonzero z component of $\vec{L} = L \, \hat{z}$.

So far, everything we have done is valid even if $V(r)$ depends on time explicitly. But usually, we are interested in a $V(r)$ that is time independent, in which case there is a conserved energy,

$$E=\frac{1}{2}m\,\dot{r}^{\,2}+\frac{1}{2}m\,r^{\,2}\,\dot{\theta}^{\,2}+V(r) \tag{20}$$

Because the angular momentum, L, is constant, we know $\dot{\theta}$ if we know r and L. Thus we can now eliminate $\dot{\theta}$ from the energy and write

$$E=\frac{1}{2}m\,\dot{r}^{\,2}+\frac{1}{2}\frac{L^2}{m\,r^2}+V(r) \tag{21}$$

This is a very important equation. As with most important equations, I hope that you will not so much memorize it as understand the logic so well that you can reproduce it instantly whenever you need it. In this case, the logic is conservation of energy and angular momentum. Among other things, (21) implies that a particle with nonzero angular momentum can never get to the origin, unless the potential goes to minus infinity as fast as $1/r^2$. The $\frac{1}{2}\frac{L^2}{m\,r^2}$ acts as a barrier (sometimes called the "angular momentum barrier" — clever, no?) that keeps it away. Thus for reasonable potentials, there is some radius of closest approach to the origin for any motion with nonzero L.

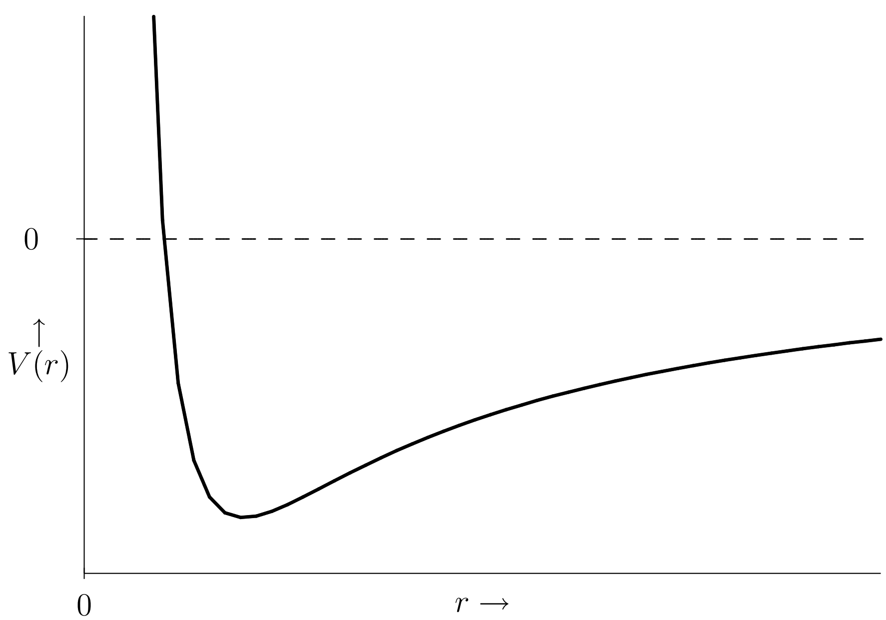
$V(\infty)=0$

We can learn a certain amount about what goes on for a particular central force law just from (21).

Consider an attractive potential that goes to zero as $r\to\infty$. For definiteness, lets consider a class of potentials

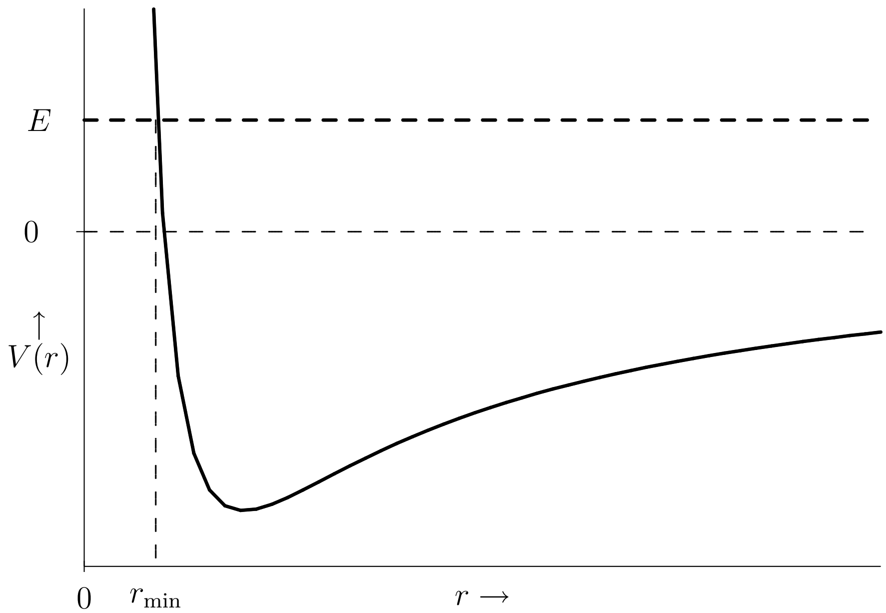
$$V(r)=-\alpha\,r^{-\beta} \tag{22}$$

for positive α and β. This is nice because it includes the gravitational potential. For potentials of this form, (21) looks more of less like this:



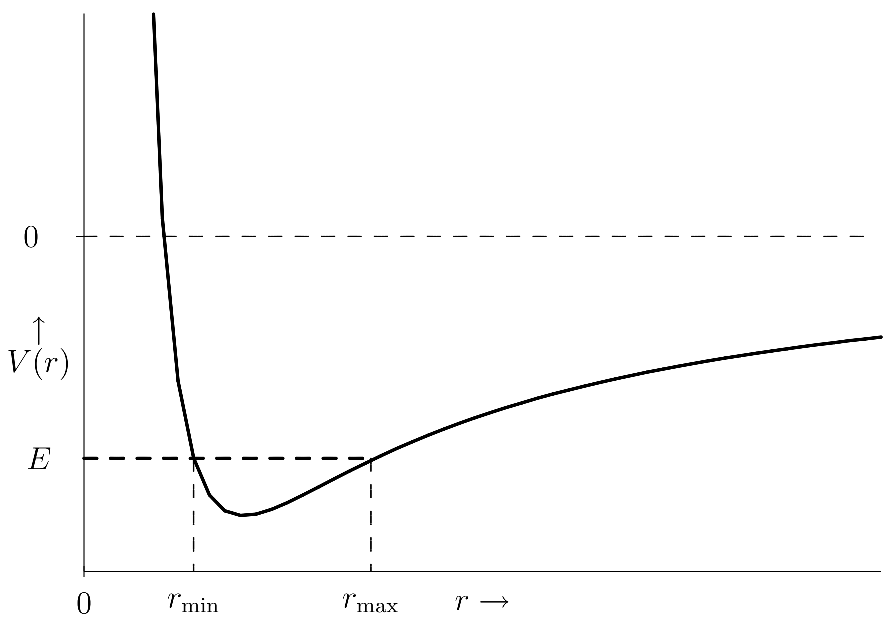
(23)

The nature of the orbit depends on the energy. If E is positive, there is a point of closest approach, but the particle then goes out to infinity with nonzero kinetic energy.



(24)

Basically, orbits with positive energy look pretty much the same so long as the potential is not really weird. But if E is negative, then things get more interesting. The particle cannot get out to $r = \infty$ so the orbiting mass must move back and forth between a minimum and maximum radius.

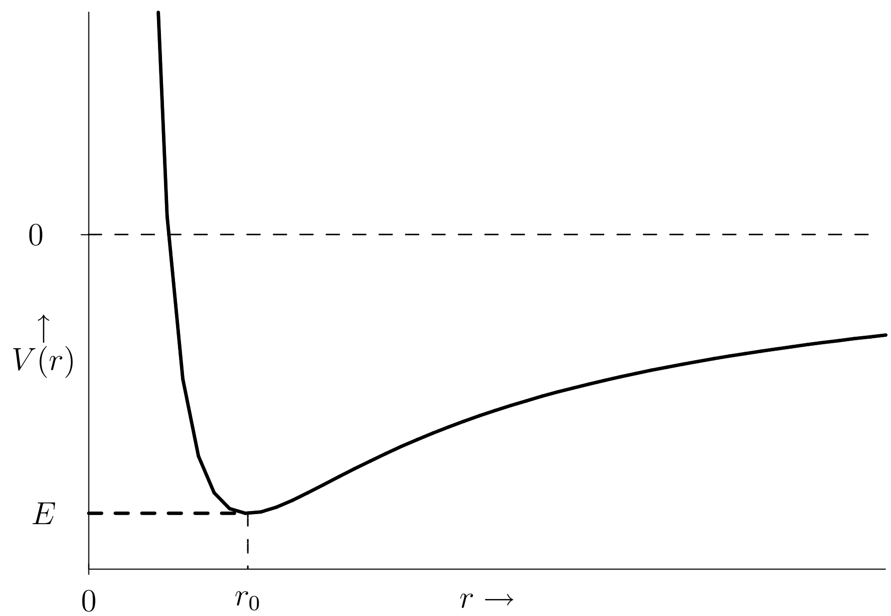


(25)

But the details of what the orbit looks likes in two dimensions is very complicated in general. For the $-1/r$ potential, the orbit closes. But that is not true for $\beta \neq 1$. We can look at these numerically in central-force.exe.

Note that except for $\beta = 1$, the point of closest approach to the origin (called the "perihelion" — the point farthest away is called the "aphelion", though this term is almost never used) moves or "precesses" as the mass orbits. The advance of the perihelion of mercury is a famous test of Einstein's theory of general relativity.

For the minimum possible energy, things get simple again. The minimum possible orbit for a given angular momentum corresponds to a circular orbit.



The parallel axis theorem

Our expressions for moment of inertia and angular momentum about an axis have so far assumed that the axis goes through the origin of the coordinate system. If instead, the axis goes through some arbitrary point R in the direction $\hat{n}$, the velocity of the point $\vec{r}_j$ is given by

$$\dot{\theta}\,\hat{n}\times(\vec{r}_j-\vec{R}) \tag{26}$$

and the moment of inertia about this axis is

$$I_{\vec{R}}=\sum_j m_j\,|\hat{n}\times(\vec{r}_j-\vec{R})|^2 \tag{27}$$

This is valid for any point $\vec{R}$, but the result can be put into a particularly simple form if $\vec{R}$ is the center of mass, defined by

$$\vec{R}=\frac{\sum_j m_j\,\vec{r}_j}{\sum_j m_j}=\frac{\sum_j m_j\,\vec{r}_j}{M} \tag{28}$$

where $M=\sum_j m_j$ is the total mass,

$$\sum_j m_j\,r_j=M\,\vec{R} \tag{29}$$

Expanding (27) gives

$$
\begin{aligned}
I_{CM} &= \sum_j m_j\,|\hat{n}\times\vec{r}_j|^2 - 2\sum_j m_j\,(\hat{n}\times\vec{r}_j)\cdot(\hat{n}\times\vec{R}) + \sum_j m_j\,|\hat{n}\times\vec{R}|^2 \\
&= \sum_j m_j\,|\hat{n}\times\vec{r}_j|^2 - 2\left(\hat{n}\times\left(\sum_j m_j\,\vec{r}_j\right)\right)\cdot(\hat{n}\times\vec{R}) + \left(\sum_j m_j\right)|\hat{n}\times\vec{R}|^2 \\
&= \sum_j m_j\,|\hat{n}\times\vec{r}_j|^2 - 2M(\hat{n}\times\vec{R})\cdot(\hat{n}\times\vec{R}) + M\,|\hat{n}\times\vec{R}|^2 \\
&= I - M\,|\hat{n}\times\vec{R}|^2 = I - M\,R^2
\end{aligned}
\tag{30}
$$

where R is the distance of the center of mass from the original axis. We can rewrite (30) as the "parallel axis theorem",

$$I = I_{CM} + M\,R^2 \tag{31}$$

which in words says that the moment of inertia about an arbitrary axis is the moment of inertia about an axis through the center of mass plus the moment of inertia of a point mass M a distance R from the axis.

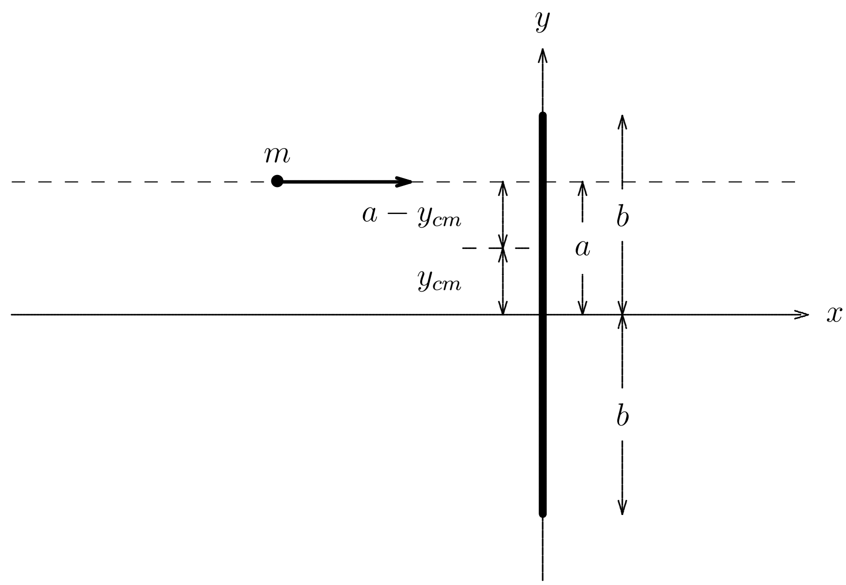
Example — Planar bodies moving in a plane

A simple situation in which we can use similar ideas in a more general way is the motion of planar rigid bodies in the plane. This sounds like a special situation, but in fact, it includes a lot of interesting situations. The angular momentum in this situation is always perpendicular to the plane. The important point here is that the configuration of a planar rigid body moving in a plane can be specified by the angle θ that gives the orientation of the rigid body in the plane, plus the position of the center of mass $\vec{r}$, where $\vec{r} = (x, y)$ is a two dimensional vector in the plane. The kinetic energy of such a body is

$$\frac{1}{2}I\,\dot{\theta}^2 + \frac{1}{2}m\,\dot{\vec{r}}^2 \tag{32}$$

where m is the mass of the body and I is the moment of inertia of the body about an axis perpendicular to the plane through its center of mass. If there are several such objects, the kinetic energy will be a sum, and the interactions will typically conserve momentum in the plane and angular momentum perpendicular to the plane.

Here is a simple example. Suppose a point mass m slides with speed v in the plane in the x direction at $y = a$ and collides and sticks to a uniform rod of mass μ of length ℓ intially at rest centered at the origin and stretched along the y axis from $y = -b$ to $y = b$, as shown below:



(33)

The center of mass of the final system has y coordinate

$$y_{cm} = \frac{1}{m+\mu}[m\,a + \mu\,0] = \frac{m\,a}{m+\mu} \tag{34}$$

After the collision, the point on the rod-mass system at $y = y_{cm}$ moves in the x direction with constant velocity

$$\frac{m\,v}{m+\mu} \tag{35}$$

To determine the rate at which the rod-mass system rotates about the center of mass, we first note that the angular momentum about the center of mass is

$$I\,\omega = -m\,v\,\frac{\mu\,a}{m+\mu} \tag{36}$$

The moment of inertia about the center of mass can be computed by adding the contribution from the point mass

$$I_m = m\,(a - y_{cm})^2 = m\,\left(\frac{\mu\,a}{m+\mu}\right)^2 \tag{37}$$

to that from the rod, which from the parallel axis theorem is

$$I_\mu = I_0 + \mu\,y_{cm}^2 = \frac{\mu\,b^2}{3} + \mu\,\left(\frac{m\,a}{m+\mu}\right)^2 \tag{38}$$

where we have used the fact that the moment of inertial of the rod about **its** center of mass is $\mu\,b^2/3$. Thus the moment of inertial of the rod-mass system about its center of mass is

$$I = m\,\left(\frac{\mu\,a}{m+\mu}\right)^2 + \frac{\mu\,b^2}{3} + \mu\,\left(\frac{m\,a}{m+\mu}\right)^2 = \frac{\mu\,b^2}{3} + \frac{m\mu\,a^2}{m+\mu} \tag{39}$$

Putting all this together, we can work out how the system moves by putting together the uniform motion of the center of mass with the rotation about the center of mass. This system is animated in the *Mathematica* notebook **rodball.nb** on the web page. The animation allows you to vary the parameters m and a, and also to make the center of mass visible, so you see more easily its uniform motion and the rotation of the system around it. I hope that you will play with this one.

Torque - fixed reference point

We believe that at the deepest level, the Lagrangian of the world is rotation invariant, and that total angular momentum is conserved (at least if the reference point is fixed), like total energy and total momentum. But a system that is not isolated from the rest of the universe may have an angular momentum that is not conserved, just as it may have a momentum or energy that is not conserved. We describe this situation in terms of the concept of torque. The angular momentum of a rigid body about a fixed reference point $\vec{r}_0$ is

$$\vec{L}_{r_0} = \sum_j (\vec{r}_j - \vec{r}_0\,) \times \vec{p}_j = \sum_j m_j\,(\vec{r}_j - \vec{r}_0\,) \times \dot{\vec{r}}_j \tag{40}$$

If we then define the torque as

$$\vec{\tau}_{r_0} = \sum_j (\vec{r}_j - \vec{r}_0) \times \vec{F}_j \tag{41}$$

Then

$$\frac{d}{dt}\vec{L}_{r_0} = \sum_j m_j\,\dot{\vec{r}}_j \times \dot{\vec{r}}_j + \sum_j m_j\,(\vec{r}_j - \vec{r}_0\,) \times \ddot{\vec{r}}_j \tag{42}$$

$$= \sum_j (\vec{r}_j - \vec{r}_0) \times \vec{F}_j = \vec{\tau}_{r_0} \tag{43}$$

The first term on the right hand side of (42) vanishes because of the antisymmetry of the cross product.

The fixed reference point $\vec{r}_0$ is completely arbitrary. This just corresponds to our freedom to choose the origin of the coordinate system. The vector $\vec{r}_j - \vec{r}_0$ describes the position of the jth mass in a coordinate system with origin $\vec{r}_0$.

It is also worth noting that the the forces in (43) need only include the "external" forces that prevent the body from being isolated and that break rotation invariance for the non-isolated system. This must be true for general reasons, and we can also see it explicitly if we make some more explicit assumptions.

The explicit calculational reason that Morin describes assumes that we can break up the force into forces between "point particles," that we all the forces involve only two particles at a time and that the force between point particles is a central force. The central force assumption is not unreasonable because the rotation invariance of the underlying theory implies that the force between two point particles in the system must be along the line between them because there is no other vector along which it could point. Then we can use Newton's third law. Thus if $\vec{F}_{ab}$ is the force on particle a from particle b,

$$\vec{F}_{ab} = -\vec{F}_{ba} \propto \vec{r}_a - \vec{r}_b \tag{44}$$

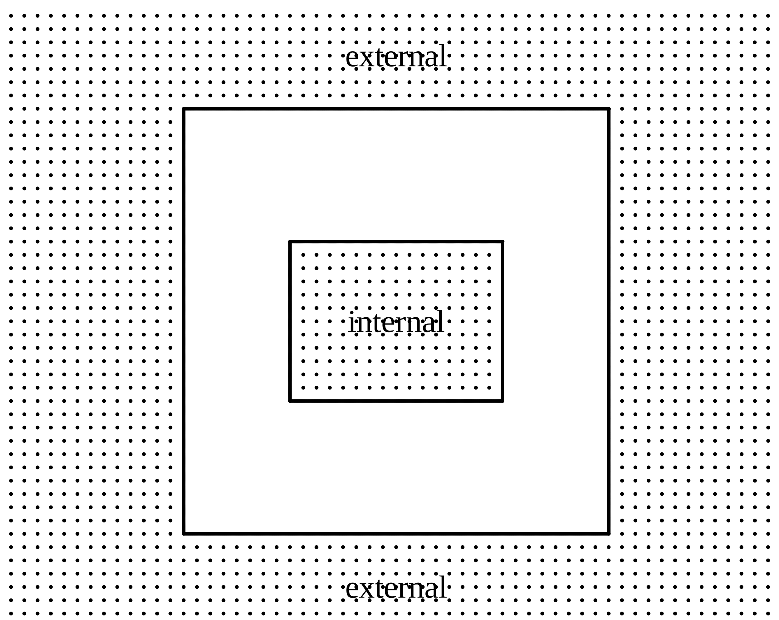
The first equality is Newton's third law, and the second comes from the assumption of central force. But then the contribution to the torque from this pair of forces vanishes because

$$(\vec{r}_a - \vec{r}_0) \times \vec{F}_{ab} + (\vec{r}_b - \vec{r}_0) \times \vec{F}_{ba} \propto (\vec{r}_a - \vec{r}_b) \times (\vec{r}_a - \vec{r}_b) = 0 \tag{45}$$

This is the argument in Morin.

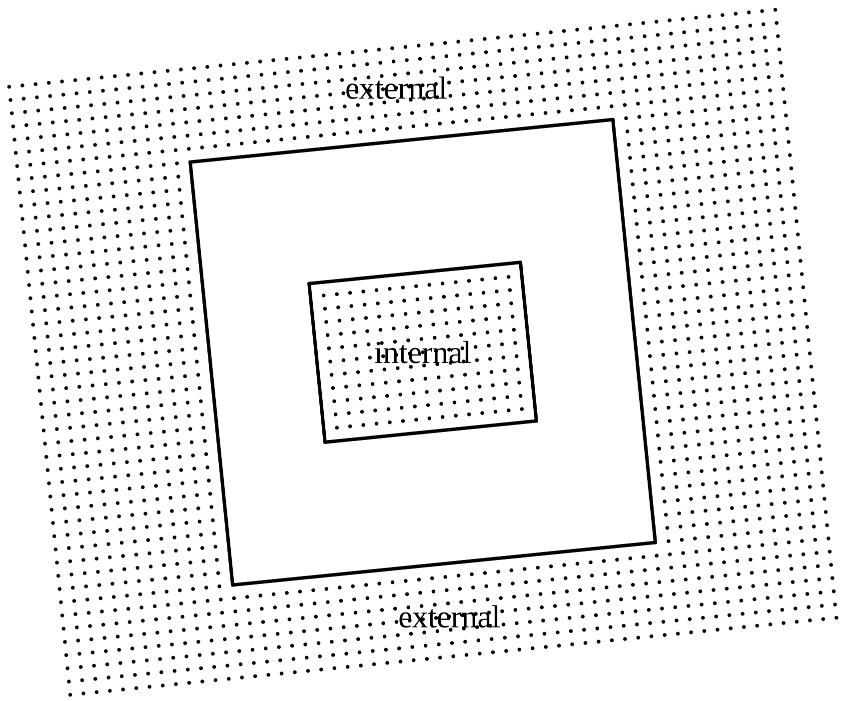
But we really don't need to make such explicit assumptions. The general reason the torque from internal forces vanishes is that, as we noted, we believe that in any really isolated system, rotation invariance holds and angular momentum is really concerved. Any violation of angular momentum conservation arises simply because we have not isolated the system. The internal forces that hold the body together do not contribute to the torque, because what "internal" means is that they don't depend on what is going on outside the system. They are invariant under rotations if as we believe the full theory of the system in isolation is invariant under rotations. The forces that hold the body together simply rotate along with the body when it is rotated. If the external forces went away, the internal forces would still be there holding the body together, but there would be no torque because the body would be isolated and rotation invariance would not be broken.

In pictures, what we are doing is separating a part of our system that we call "internal."

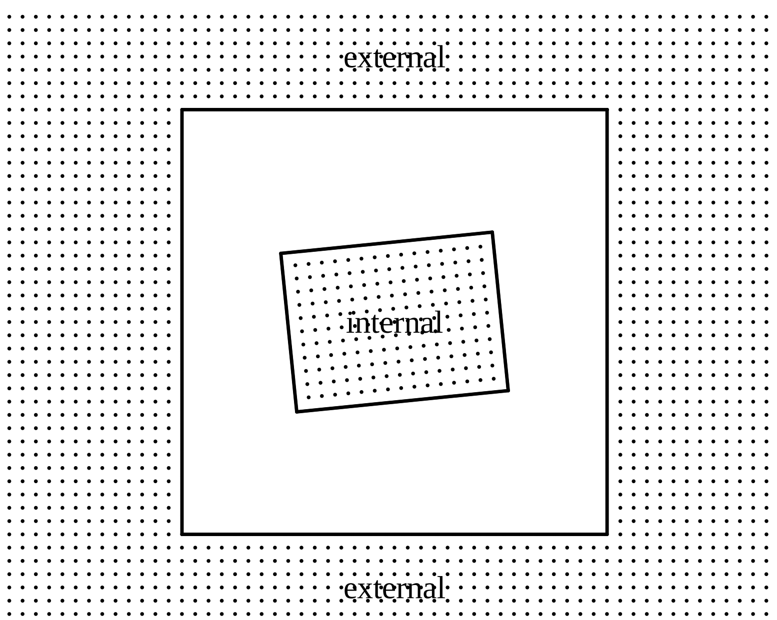


If we rotate the internal and the external parts of the system together, that has no effect on the

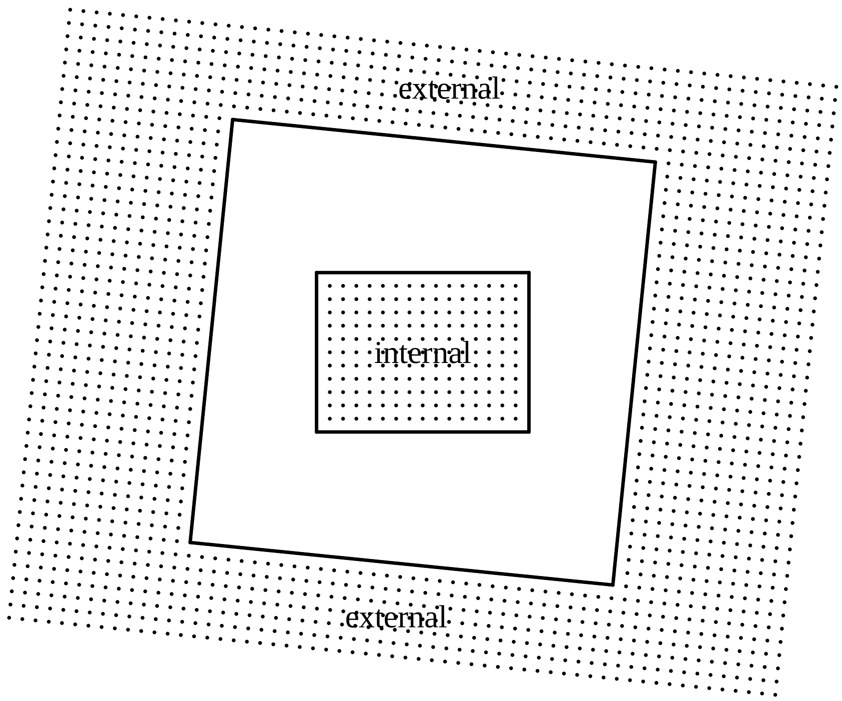
Lagrangian. The energy doesn't change. There is no torque.



But if we rotate the internal parts of the system, while keeping the external parts of the system fixed, then the energy can change and there be a torque. But this depends only on the relative orientation of what we have called internal and external. We haven't changed the internal system itself.



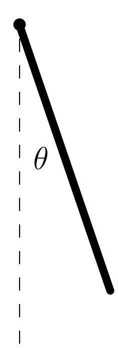
In fact, we could get the same torque by keeping the internal system completely fixed, and rotating the external system in the opposite direction because this doesn't change the **relative orientation** of internal and external, and because of overall rotation invariance, that is all that matters. Obviously, in this case we have not changed the internal system at all, so there is no purely internal contribution to the torque.



Example — oscillations of a hanging rod

Here's a simple example of using torque and moment of inertia — consider the small oscillations of a system of a solid rod of length ℓ with uniform linear mass density ρ pivoted at the top and allowed to rotate in a vertical plane in the earth's gravitational field, as shown (with the rod constrained to rotate in the plane of the paper about the pivot at the top):

(46)



We can take θ to be the angle of the rod from the vertical. The moment of inertia I is given by

$$I = \sum_j m_j \, |\hat{n} \times \vec{r}_j|^2 \tag{47}$$

To finish the job we must actually compute the moment of inertia I. We can write the "sum" in (47) as

$$I = \int_0^\ell \rho\, x^2 = \frac{1}{3}\rho\, \ell^3\, dx \tag{48}$$

Thus the angular momentum is (out of the paper)

$$L = I\dot{\theta} = \frac{1}{3}\rho\, \ell^3\, \dot{\theta} \tag{49}$$

To find the torque, we can use the fact that we can assume that the graviational force from the uniform field acts on the center of mass. This is because torque is linear in the position, so the integral would do to find the torque by adding up the contribution of all the little masses in the rod is the same integral we would do to find the center of mass.

$$\int \vec{r} \times d\vec{F} = \int \vec{r} \times (-g\hat{z})\, dm = \left(\int \vec{r}\, dm\right) \times (-g\hat{z}) = M\vec{R} \times (-g\hat{z}) = \vec{R} \times \vec{F} \tag{50}$$

This is

$$\tau = -\rho\ell g\, (\ell/2) \sin\theta = -\frac{1}{2}\rho\ell^2 g \sin\theta \tag{51}$$

again out of the paper. Thus using the fact the torque is rate of change of angular momentum, we can write

$$\frac{d}{dt} I\dot{\theta} = I\ddot{\theta} = \frac{1}{3}\rho\, \ell^3\, \ddot{\theta} = -\frac{1}{2}\rho\ell^2 g \sin\theta \tag{52}$$

or

$$\ddot{\theta} = -\frac{3}{2}\frac{g}{\ell} \sin\theta \tag{53}$$

For small θ, this is a harmonic oscillator with angular frequency

$$\omega = \sqrt{\frac{3g}{2\ell}} \tag{54}$$

Notice, however, that the result is quite different from what we would get for a simple pendulum with a mass at the center of mass of the rod. That would have angular frequency

$$\omega = \sqrt{\frac{2g}{\ell}} \tag{55}$$

The effect of the moment of inertia is to add some extra inertia to the system which decreases the angular frequency and increases the period of oscillation.

We could do this somewhat more easily using the Lagrangian, without mentioning torque. The kinetic energy is

$$\frac{1}{2} I\dot{\theta}^2 \tag{56}$$

The potential energy can be taken to be

$$\rho\ell g\, \frac{\ell}{2}\, (1 - \cos\theta) \tag{57}$$

(the constant 1 is not important, but I guess it is nice to have the energy defined to be 0 at $\theta = 0$). Then the Lagrangian of the system is

$$\mathcal{L} = \frac{\rho\ell^3}{6}\,\dot\theta^2 - \frac{\rho\ell^2}{2}\,(1-\cos\theta) \tag{58}$$

and the Euler-Lagrange equation is

$$\frac{\rho\ell^3}{3}\,\ddot\theta = -\frac{\rho g\ell^2}{2}\,\sin\theta \tag{59}$$

Torque - moving reference point

Things are slightly more complicated when the reference point is moving - when $\dot{\vec r}_0 \neq 0$. But only a little, so long as we look at things in a reference frame that is moving along with the reference point. The first thing to note is that this absolutely has to work the same way if

$$\ddot{\vec r}_0 = 0 \tag{60}$$

because in this case the frame in which reference point is fixed is moving with constant velocity and is therefore an inertial frame, just as good as the one we started in. In the moving reference frame, positions and velocities are different, but forces and accelerations are all the same, and Newton's laws work the same way.

The angular momentum in the moving frame in terms of the original $\vec r$ looks like

$$\vec L_{r_0} = \sum_j m_j\,(\vec r_j - \vec r_0\,) \times (\dot{\vec r}_j - \dot{\vec r}_0\,) \tag{61}$$

We get the second terms in the parentheses because we are measuring the position and velocity in the moving frame and have to subtract the new reference point.

If the reference point is accelerating, $\ddot{\vec r}_0 \neq 0$, this is a dangerous thing to do, because Newton's laws don't work quite the same way in an accelerating frame, but we will see how that works in a moment. However, nothing prevents us from considering $\vec L_{r_0}$ defined by (61) in this way, so let's try it. Now we can just differentiate and see what happens.

$$\frac{d}{dt}\vec L_{r_0} = \sum_j m_j\,(\dot{\vec r}_j - \dot{\vec r}_0\,) \times (\dot{\vec r}_j - \dot{\vec r}_0\,) + \sum_j m_j\,(\vec r_j - \vec r_0\,) \times (\ddot{\vec r}_j - \ddot{\vec r}_0\,) \tag{62}$$

Again the first term on the right hand side vanishes because of the antisymmetry of the cross product and this looks just like (42) except for the $\ddot{\vec r}_0$ term.

$$\frac{d}{dt}\vec L_{r_0} = \vec\tau_{r_0} - \sum_j m_j\,(\vec r_j - \vec r_0\,) \times \ddot{\vec r}_0 \tag{63}$$

There are a couple of important situations in which we can simply ignore the term proportional to $\ddot{\vec r}_0$. One, as we expected, is when the second derivative vanishes,

$$\ddot{\vec r}_0 = 0 \text{ — that is if } \vec r_0 \text{ is not accelerating.} \tag{64}$$

The other is more interesting. We can rewrite the $\ddot{\vec{r}}_0$ term using the definition of the center of mass as

$$-\sum_j m_j\,(\vec{r}_j - \vec{r}_0\,) \times \ddot{\vec{r}}_0 \;= -M\,(\vec{R} - \vec{r}_0\,) \times \ddot{\vec{r}}_0 \tag{65}$$

Evidently this vanishes if

$$\vec{r}_0 = \vec{R} \text{ — that is if } \vec{r}_0 \text{ is the center of mass;} \tag{66}$$

or if

$$(\vec{r}_0 - \vec{R}) \text{ is parallel to } \ddot{\vec{r}}_0 \text{ so the cross product vanishes.} \tag{67}$$

Then

$$\frac{d}{dt}\vec{L}_{r_0} = \vec{\tau}_{r_0} \tag{68}$$

where the torque τ_{r_0}

$$\vec{\tau}_{r_0} = \sum_j (\vec{r}_j - \vec{r}_0) \times \vec{F}_j \tag{69}$$

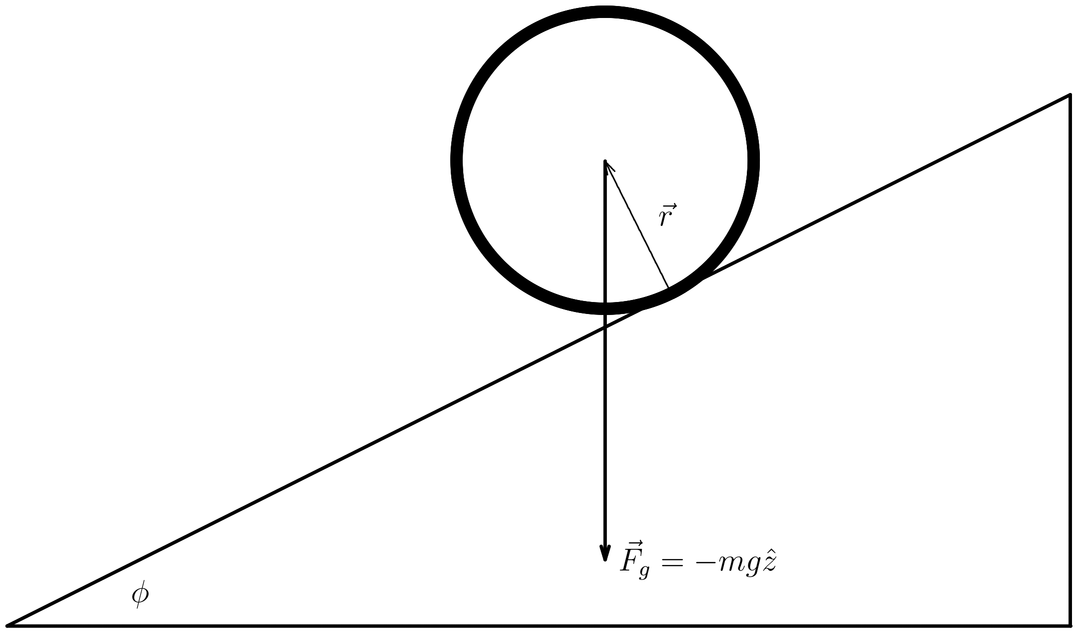
is computed in terms of the forces in the lab frame.

Thus we can continue to use $\vec{\tau} = d\vec{L}/dt$ if we take the reference point to be the center of mass **even if the center of mass is accelerating!** The best way of understanding this is something we will discuss in more detail next month. We can actually use Newtonian physics in an accelerating reference frame if we also introduce what is called a "fictitious force" to make up for the effect of the acceleration. Next month, we will argue that we can simply put the $\ddot{\vec{r}}_0$ on the other side of the $\vec{F} = m\vec{a}$ equation and associate it with such an "fictitious force." This force behaves exactly like a uniform gravitational force (indeed, Einstein would say that it is not really fictitious - it is entirely equivalent to a gravitational force - this is the starting point for general relativity - but we won't go there right now). But a uniform gravitational force just acts on the center of mass, and it produces no torque about the center of mass. That is why this works.

Thus if we take the reference point $\vec{r}_0$ to be the center of mass $\vec{R}$, the rate of change of angular momentum about $\vec{r}_0$ is equal to the torque about $\vec{r}_0$, even if $\vec{r}_0$ is moving. This is a good thing, so I will assume that the reference point is the center of mass unless explicitly stated. From now on, I will drop the subscript R.

Example — Rolling ring

Here is another example — a ring of mass m and radius r rolling without slipping down an inclined plane in a gravitational field.



(70)

This is a system with one degree of freedom. We can specify the position of the ring either by the distance ℓ that the ring rolls, or by the angle θ through which the ring rolls, and these are related by

$$\ell = r\theta \tag{71}$$

Let's do this three ways. First let's consider the torque and angular momentum about the reference point defined by the point where the ring touches the plane. The nice thing about this reference point is that the torque on the ring is independent of the force that the plane exerts on the ring. It comes entirely from the gravitational force, which as usual we can take to act at the center of mass. This point is instantaneously at rest, and the motion of the ring at the instant it is at rest is a rotation about this reference point Then the torque is

$$mgr\sin\phi \tag{72}$$

The moment of inertia is given by the parallel axis theoren as

$$I = mr^2 + I_{CM} = 2mr^2 \tag{73}$$

because the mass is all a distance r away from the center of mass. Thus we have

$$\dot{L} = I\ddot{\theta} = 2mr^2\ddot{\theta} = mgr\sin\phi \tag{74}$$

or

$$\ddot{\theta} = \frac{g}{2r}\sin\phi \tag{75}$$

This describes uniform angular acceleration. In terms of $r\theta$, the distance down the inclined plane, it is uniform acceleration with $g\sin\phi/2$. Note that there is an extra factor of 1/2 compared with mass sliding down a frictionless plane with the same angle. This is the effect of the extra inertia associated with the moment of inertia.

I glossed over an important point here, and let me now emphasize it. Our reference point in this analysis is not moving. We are looking at the system at a particular time, t, and the reference point is the fixed point where the ring touches the plane at that one time — not the moving point where the ring touches the plane as a function of time. The way this shows up in the calculation is that we computed the angular momentum in the lab frame. If we had been using a moving reference frame, we would have had to measure the velocities of the parts of the system with respect to that moving frame. That would have been more complicated - and also useless in this case, because the point at which the ring touches the plane is accelerating as a function of time.

Now let's do the problem using the Lagrangian. As the ring rolls without slipping, its center of mass moves down the plane, and it also spins. Thus its kinetic energy is a sum of a linear and a rotational term. If we describe the configuration of the system by the angle θ that the ring rolls, the linear distance it moves is $r\,\theta$. Thus the kinetic energy is

$$\frac{1}{2}m\,(r\,\dot{\theta})^2 + \frac{1}{2}I_{CM}\,\dot{\theta}^2 \tag{76}$$

and I_{CM} in this case is just $m\,r^2$ because all the mass of the ring is a distance r from the center of mass. The gravitational potential is

$$-\,mgr\theta\sin\phi \tag{77}$$

Thus the Lagrangian is

$$m\,r^2\,\dot{\theta}^2 + mgr\theta\sin\phi \tag{78}$$

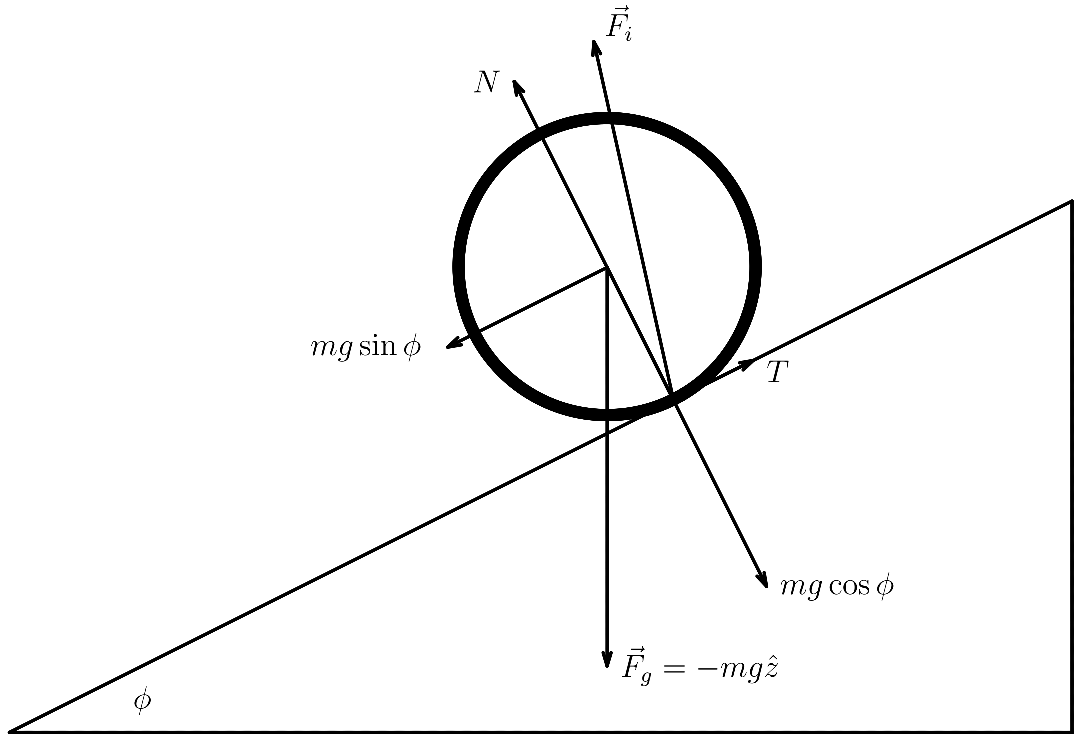
and the Euler-Lagrange equation is

$$2m\,r^2\,\ddot{\theta} = mgr\sin\phi \tag{79}$$

which again gives

$$\ddot{\theta} = \frac{g}{2r}\sin\phi \tag{80}$$

Another way to do the problem is in term of torque and rate of change of angular momentum about the center of mass. But to do it this way, we need to consider the force $\vec{F}_i$ that the inclined plane exerts on the ring, as shown below.



(81)

The normal component of $\vec{F}_i$ (labeled by N in the diagram) just cancels the normal component of the gravitational force, $mg\cos\phi$, because the ring doesn't accelerate off the plane. The tangential component of $\vec{F}_i$ (labeled by T in the diagram) is the frictional force that keeps the ring from sliding on the incline. Now the acceleration of the ring down the incline is determined by the difference between the tangential component of the gravitational force and T,

$$ma = mg\sin\phi - T\,. \tag{82}$$

But the angular acceleration of the ring is governed by the moment of inertia, $I = mr^2$, and the torque about the center of mass, which is out of the plane of the paper with magnitude rT,

$$mr^2\ddot{\theta} = rT \tag{83}$$

Because the ring is rolling without slipping

$$a = r\ddot{\theta} \tag{84}$$

so (82) becomes

$$mr\ddot{\theta} = mg\sin\phi - T\,. \tag{85}$$

Multiplying (85) by r and adding it to (83), we can cancel the unknown T and find

$$2mr^2\ddot{\theta} = mgr\sin\phi \tag{86}$$

which agrees with (79) as it should.

As we have seen before in other contexts, the Lagrangian technique simplifies things. It allows us to solve the problem without introducing and then eliminating the frictional force, T. But we were able to do it just as easily by choosing the reference point appropriately

lecture 18

Topics:

- Where are we now?
- Impulse and elastic collisions
- Rigid bodies are weird
- The angular velocity vector
- Impulse and rigid bodies
- The moment of inertia tensor
- An impulsive demo

Where are we now?

We began our discussion of rigid body rotations by discussing the simple case of rotations about a fixed axis. Today, I will spend a little time discussing in general the important ideas of the angular velocity vector and the reference point. I am also going to discuss a nice example of motion in a plane, which introduces the idea of impulse that we will use to explore more complicated situations.

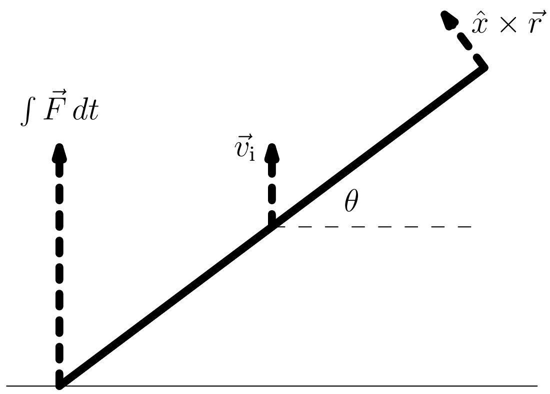
Impulse and elastic collisions

The animation in the *Mathematica* notebook rodbounce.nb shows a rigid rod in the y-z plane in a gravitational field bouncing completely elastically on a frictionless surface. This problem is a nice example of the use of impulsive forces and torques to solve problems. I will start by analyzing this example to show you how the animation was produced.

We will assume that the rod is initially either not rotating at all (this is what is shown in the animation) or rotating in a vertical plane. If so, the motion stays in the same vertical plane, as long as the rod is perfectly symmetrical and the plane on which it bounces is perfectly flat. Thus we can analyze it without worrying about the full three dimensional complexity of angular momentum. Then we can just choose our coordinate system so that $\hat{z}$ is vertical and the rod is bouncing in the y-z plane, as we assumed.

Except when the rod is actually in contact with the frictionless surface, the motion is extremely simple. The rod rotates with some fixed angular velocity ω_{i} — the subscript i is for "initial" (the axis $\hat{x}$ is out of the plane in the x direction) and the center of mass, in the center of the rod rises and falls in the constant gravitational field, so that when one end of the rod hits the frictionless surface, the center of mass is moving with some velocity v_{i} in the vertical direction ($\vec{v} = v\,\hat{z}$), which will usually be negative, but not always, because the rotation of the rod may cause a collision even if the center of mass is rising. When a collision occurs, we get a new velocity and angular velocity, v_{f} and ω_{f} (subscript f for "final"). Our job is to calculate v_{f} and ω_{f} in terms of v_{i} and ω_{i} — then we can follow the system until the next collision and do it again, and so on until we get tired. Or better still, we can simply program it into the animation and watch the pretty bouncing rod until we get mesmerized.

Suppose that the rod has mass m and length 2ℓ. Suppose further that the collision occurs with the rod at an angle θ (between 0 and π) from the horizontal, as shown below.



(1)

Also shown is the translational velocity of the center of mass and the rotational motion, $\omega\hat{x} \times \vec{r}$, which gets added on to the motion of the center of mass to produce the full motion of the end of the rod.

During the collision, there is a force, $\vec{F}$, on the rod from the frictionless surface. Because the surface is frictionless, the force is purely vertical, $\vec{F} = F\,\hat{z}$. Now because the bounce happens very quickly, we can ignore the motion of the rod while the bounce is taking place. Then all that matters is the integral of the force over the period of the bounce,

$$\int_{\rm bounce} dt\,\vec{F} = \hat{z}\int_{\rm bounce} dt\,F \tag{2}$$

This is called the "impulse." Now the point is that the impulse does double duty. 1 — It changes the linear momentum of the center of mass. Because the force is the rate of change of the momentum, the impulse is the total change in the momentum:

$$m\,(v_{\rm f} - v_{\rm i}) = \int_{\rm bounce} dt\,\frac{dp}{dt} = \int_{\rm bounce} dt\,F \tag{3}$$

2 — It also changes the angular momentum about the center of mass. Because the torque is the rate of change of angular momentum, the cross product of the lever arm with the impulse is the total change in the angular momentum. The rod has length 2ℓ, so this looks like

$$I\,(\omega_{\rm f} - \omega_{\rm i}) = \int_{\rm bounce} dt\,(\vec{r}\times\vec{F})_x = \left(\vec{r}\times\int_{\rm bounce} dt\,\vec{F}\right)_x = -\ell\cos\theta\int_{\rm bounce} dt\,F \tag{4}$$

Let me emphasize again the key step here. Because we have assumed that the bounce takes place very quickly, we can ignore the motion of the rod while the bounce is taking place. That allows us to take the $\vec{r}$ out of the integral in (4). This makes the problem doable because the change in v and the change in ω are related —

$$\ell\cos\theta\,m\,(v_{\rm f} - v_{\rm i}) + I\,(\omega_{\rm f} - \omega_{\rm i}) = 0 \tag{5}$$

In addition to (5), we know that energy is conserved. The energy is the kinetic energy in motion of the center of mass and rotation. Thus conservation of energy is

$$\frac{1}{2}\,m\,v_{\rm i}^2 + \frac{1}{2}\,I\,\omega_{\rm i}^2 = \frac{1}{2}\,m\,v_{\rm f}^2 + \frac{1}{2}\,I\,\omega_{\rm f}^2 \tag{6}$$

At this point, we could plug this into Maple or Mathematica and ask the computer to solve for $v_{\rm f}$ and $\omega_{\rm f}$ in terms of $v_{\rm i}$ and $\omega_{\rm i}$. But there is a useful trick involved in doing it by hand, so let's go on a while. First write energy conservation as

$$m\left(v_{\rm f}^2 - v_{\rm i}^2\right) + I\left(\omega_{\rm f}^2 - \omega_{\rm i}^2\right) = 0 \tag{7}$$

Now we can factor this

$$m\left(v_{\rm f} - v_{\rm i}\right)\left(v_{\rm f} + v_{\rm i}\right) + I\left(\omega_{\rm f} - \omega_{\rm i}\right)\left(\omega_{\rm f} + \omega_{\rm i}\right) = 0 \tag{8}$$

and now use (5) to write this as

$$\begin{aligned} &m\left(v_{\rm f} - v_{\rm i}\right)\left(v_{\rm f} + v_{\rm i}\right) - \ell\cos\theta\, m\left(v_{\rm f} - v_{\rm i}\right)\left(\omega_{\rm f} + \omega_{\rm i}\right) \\ &= m\left(v_{\rm f} - v_{\rm i}\right)\Big(\left(v_{\rm f} + v_{\rm i}\right) - \ell\cos\theta\left(\omega_{\rm f} + \omega_{\rm i}\right)\Big) = 0 \end{aligned} \tag{9}$$

The trick here is pretty general. We know there is a solution to the twin equations (5) and (6) of the form $v_{\rm f} = v_{\rm i}$ and $\omega_{\rm f} = \omega_{\rm i}$ because this satisfies (5) and if nothing changes, energy is conserved. We have just written (6) so that this solution is manifest. Of course, we are not interested in the case where nothing changes because this is not what happens in the collision. There must always be some force on the rod during the collision, so we always get a non-trivial change in v and ω. But writing (6) this way allows us to eliminate the trivial solution and makes it easier to find the interesting one. Thus for the physical solution we are interested in, we must have

$$\left(v_{\rm f} + v_{\rm i}\right) - \ell\cos\theta\left(\omega_{\rm f} + \omega_{\rm i}\right) = 0 \tag{10}$$

This is now another linear equation for $v_{\rm f}$ and $\omega_{\rm f}$, so we can easily solve (5) and (10), and the result is

$$v_{\rm f} = \frac{\left(m\ell^2\cos^2\theta - I\right)v_{\rm i} + 2I\ell\cos\theta\,\omega_{\rm i}}{m\ell^2\cos^2\theta + I} \qquad \omega_{\rm f} = \frac{2m\ell\cos\theta\, v_{\rm i} - \left(m\ell^2\cos^2\theta - I\right)\omega_{\rm i}}{m\ell^2\cos^2\theta + I} \tag{11}$$

It is this that we have used to construct the animation.

Nothing we have done depends on the precise value of I. For a solid rod, $I = m\,\ell^2/3$, but we don't have to look only at that case. It is interesting to look at this for various Is. One very interesting limit is $I = m\,\ell^2$, which corresponds to a dumbbell, with the masses at the ends of a light rod. This is animated in the *Mathematica* notebook dumbbellbounce.nb This actually allows us to use our physical intuition to get a nice check of (11). Suppose that for a dumbbell, $\cos\theta$ is close to zero when the left mass hits the surface. Because the force of the rod on the masses is nearly horizontal in this case, it has very little effect of the motion of the two masses. Thus we expect the left mass to bounce and simply reverse its velocity, and the right mass to just keep going. Now for $\theta \approx 0$, the motion of the masses is nearly in the vertical direction and the vertical components are approximately

$$v_{\rm left} \approx v - \ell\,\omega \qquad\qquad v_{\rm right} \approx v + \ell\,\omega \tag{12}$$

Thus to get $v_{\rm left}$ to change sign while $v_{\rm right}$ to remain unchanged, we want the bounce to approximately interchange v and $\ell\,\omega$,

$$v_{\rm f} \to \ell\,\omega_{\rm i} \qquad\qquad \ell\,\omega_{\rm f} \to v_{\rm i} \tag{13}$$

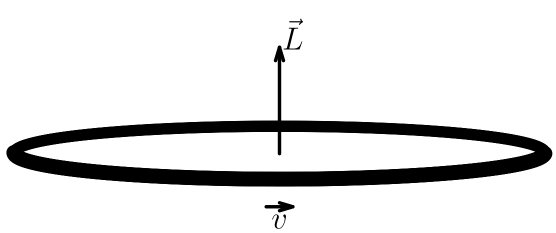
It is easy to see from (11) that this is what happens.

In the case of the rigid rod, with smaller moment of inertia, the other end of the rod actually moves faster after such a nearly horizontal bound. Again a limit may make clearer what is going on. The limit $I \to 0$ corresponds to a mass in the center of the light rod. In this case, for $\theta \approx 0$, we expect the center of mass to keep moving, $v_{\rm f} \approx v_{\rm i}$, which again accords with (11).

Meanwhile, notice that when the rod is rotating a lot, it doesn't go up as far — this is because more of the energy is stored in rotational kinetic energy and there is less in center of mass motion after the collision.

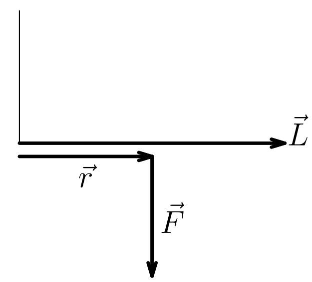
Rigid bodies are weird

So what's the big deal about torque and angular momentum. Surely, this is just like force and momentum. You push something and it moves (or accelerates, at least). You twist something and it turns. But torque equals rate of change of angular momentum implies some pretty remarkable things. When you twist a spinning rigid body carrying a large angular momentum, the twist does very counter-intuitive things because the twist does not directly change the orientation of the body. Instead, what a torque does is to change the direction of the angular momentum. And the direction of the angular momentum is tied not to the orientation of the body, which is constantly changing, but to the orientation of the rotation axis. A gyroscope is the most familiar example of this. A torque that one would naively think would cause the body to fall instead causes it to precess. This is very familiar, but it is worth seeing over and over again. Here it is for a simple top. Take a bicycle wheel and weight the rim with lead. Get it spinning with speed v. The angular momentum is then approximately mvr where r is the radius of the wheel and m is the total mass. This would be exactly right if all of the mass were concentrated in the rim at radius r.



(14)

Now if we apply a torque to the handle, which is the axis of rotation, strange things happen because we are actually changing the angular momentum. Precession is one example.



(15)

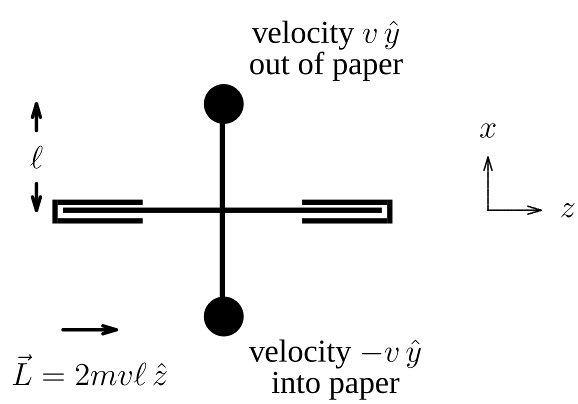
The torque in the diagram is into the paper. Thus the change in the angular momentum is into the paper. But the only way that can happen is if the direction of the angular momentum changes — and the orientation of the handle must go with the angular momentum — so the system precesses. You see that precession is easy to explain in terms of torque and angular momentum, but perhaps not so easy to understand in your bones.

Here is another situation which is basically the same, but which I find even stranger. If I stand on a turntable with the wheel axis horizontal, and I try to twist the handle so that the angular momentum of the wheel points slightly down, to conserve angular momentum, I will have to start spinning in the counterclockwise direction, and develop angular momentum upwards. This is pretty weird, because it means that by trying to produce a torque in one direction (horizontal) I have actually produced a torque in the vertical direction.

This is so strange that it seems magical, even though we have all seen it many times. What is really going on here??????

I find this sufficiently strange that I want to show you how it comes about in a particular very simple case.

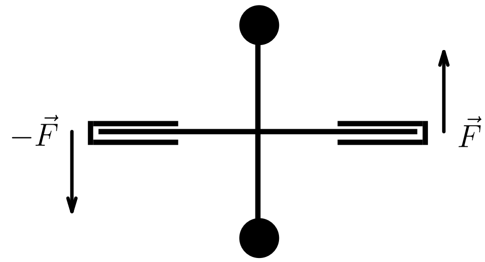
Consider a light rigid frame of crossed bars with weights of equal mass m on two of the opposite ends. The two masses form a dumbbell rotating in the x-y plane. The cross piece is supported in two frictionless sleeves that allow the system to rotate, but can be used to supply a torque. This is shown below in the x-z plane:



(16)

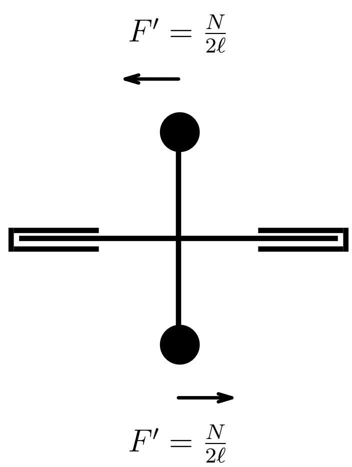
If at some time $t = 0$, the upper mass at $x = \ell$ is moving in the $+y$ direction with speed v, and the lower one at $x = -\ell$ in the $-y$ direction with the same speed, then the angular momentum of the system is $2mv\ell\,\hat{z}$.

Now suppose that at time $t = 0$, we apply a large torque N in the $+y$ direction for a very short time Δt, short enough that we can neglect the motion of the masses during the time. We actually supply this torque by twisting the frictionless sleeve,



(17)

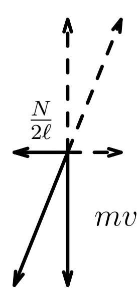
but the effect is the same as applying the forces to the masses (because of the rigidity of the frame)



(18)

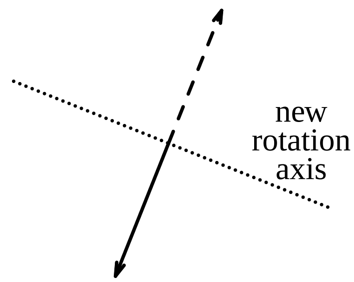
The magnitude of the force is fixed by the value of the torque.

Now the torque in (18) changes the direction of the motion of the two masses. The change in the momentum is $F'\,\Delta t$. From the top, in the y-z plane, the resulting momentum looks like this (with the momenta of the bottom mass dashed and the masses not shown)



(19)

Now as the masses move in their new direction, they drag the rest of the rigid body along with them, establishing a new axis of rotation, shown as the dotted line below.



(20)

Thus the system must precess.

Of course, this all really happens at once, but it makes it easier for me to understand what is going on in precession if I disarticulate things and think first about a very quick application of torque changing the direction of the rotating masses, and then the rigid body forces that hold the system together pulling the rest of the system along with them. Over the next couple of weeks (and starting later today if I have time), I am going to use this trick in more complicated ways to ease our way into the physics of rigid body rotations.

The angular velocity vector

We have seen that the velocity of point $\vec{r}_j$ on a rigid body rotating with angular velocity ω about an axis $\hat{n}$ through a point $\vec{r}_0$ is

$$\omega\,\hat{n} \times \left(\vec{r}_j - \vec{r}_0\right) \tag{21}$$

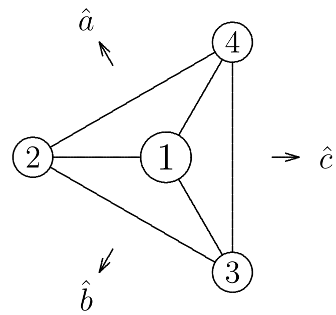
The important point is that ω and $\hat{n}$ always appear in combination, as the product $\omega\,\hat{n}$. This is a vector with direction $\hat{n}$ and magnitude ω. It is called the "angular velocity vector"

$$\vec{\omega} \equiv \omega\,\hat{n} \tag{22}$$

Now here is a very important fact. Like any vector, $\vec{\omega}$ can be taken apart into components. Angular velocity vectors can be added and subtracted. This may not sound very remarkable, but in fact, ordinary rotations do not work this way. Unlike vector coordinates like the coordinates of the center of mass, the quantities that describe the orientation of a rigid body are not vectors. They are angles and they do not form a linear space. You can't add them. The reason is simply that the structure of rotations is more complicated than the structure of translations. The order in which rotations are done matters to the final configuration. Because order doesn't matter when you add numbers or vectors, that means that rotations cannot simply add. They must compose in some more complicated way.

Here's a simple example. One good way to specify the orientation of a rigid body is by specifying a reference orientation and specifying the rotation required to get from the reference orientation to the actual orientation. A rotation, in turn, can be specified by giving the axis of rotation, and the magnitude of the rotation in radians. If you put together two rotations about the same axis, the magnitudes just add. But the trouble is that if the axes are different, the combination of the two rotations is a rotation about some new axis, by some magnitude that is a complicated function of the angles and axes. Both the axis and the magnitude of the combined rotation depend on which of the component rotations is performed first. This is what makes rotation of rigid bodies such a complicated and interesting subject.

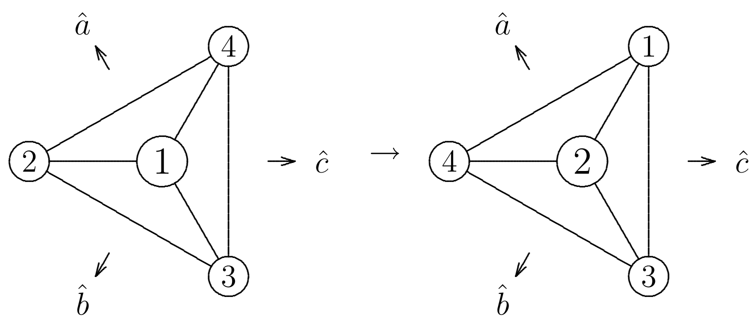
For example, consider a regular tetrahedon with the four vertices labeled by different numbers, 1-4. Looking down on this tetrahedron, it might look like this:



(23)

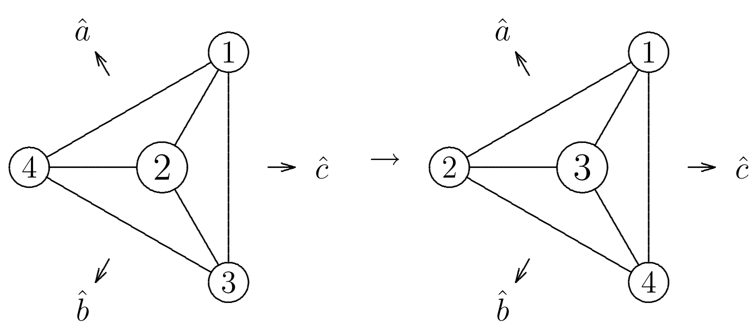
where $\hat{a}$, $\hat{b}$ and $\hat{c}$ represent axes through the center of the tetrahedron (each of these vectors is coming slightly out of the plane of the paper).

If I do a rotation by $2\pi/3$ about the $\hat{a}$ axis, the tetrahedron rotates into itself with the numbers on the vertices rotated as $1 \rightarrow 2 \rightarrow 4 \rightarrow 1$, as shown:



(24)

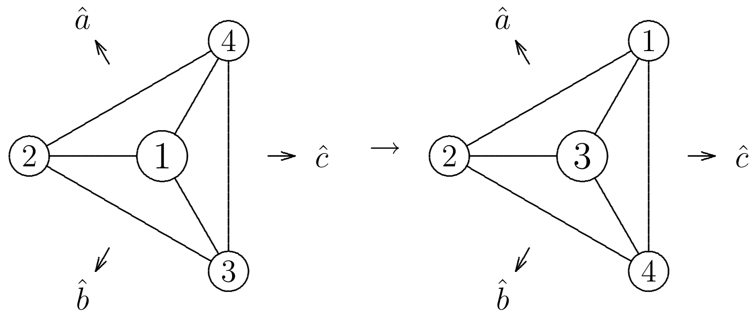
If I now do a rotation by $2\pi/3$ about the $\hat{b}$ axis, the tetrahedron rotates into itself with the numbers on the vertices rotated as $2 \rightarrow 3 \rightarrow 4 \rightarrow 2$, as shown:



(25)

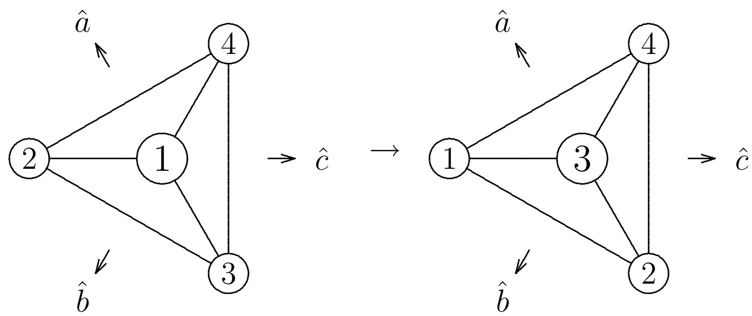
The result of these two rotations is equivalent to a single rotation by $4\pi/3$ (or $-2\pi/3$) around the

$\hat{c}$ axis, as shown below



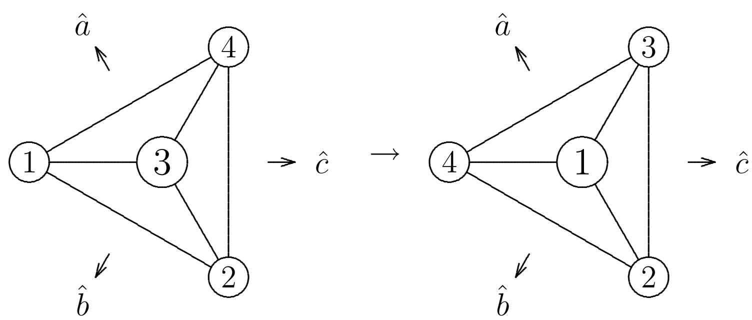
(26)

On the other hand, if I do the same two rotations in the opposite order, something different happens. If I first do the rotation by $2\pi/3$ about the $\hat{b}$ axis, the tetrahedron rotates into itself with the numbers on the vertices rotated as $1 \to 3 \to 2 \to 1$, as shown:



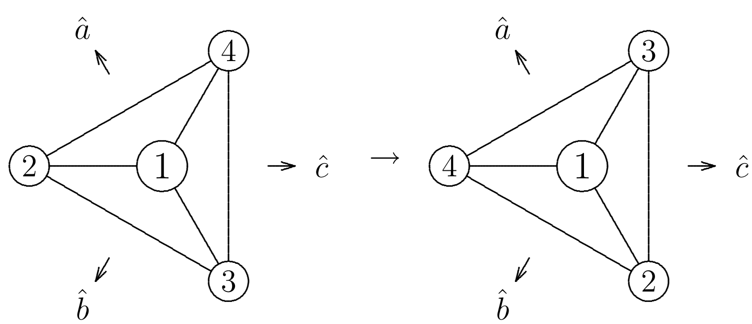
(27)

If I now do the rotation by $2\pi/3$ about the $\hat{a}$ axis, the tetrahedron rotates into itself with the numbers on the vertices rotated as $1 \to 4 \to 3 \to 1$, as shown:



(28)

The result of these two rotations is equivalent to a single rotation by $4\pi/3$ (or $-2\pi/3$) around an axis pointed down into the plane, as shown below



(29)

So you see that for finite rotations, the order of the rotations makes a difference. There is no way that you can simply add the coordinates of vectors to get these results. Finite rotations are not vectors!

The reason that angular velocities are simpler is that they really only refer to infinitesimal rotations — ω is $d\theta/dt$. And infinitesimal rotations can be added without causing confusion. Technically, the reason that infinitesimal rotations are can be added like ordinary vectors is that the depedence on the order of two infinitesimal rotations is proportional to the product of the two infinitesimal angles, and can thus be ignored. Thus $d\vec{\omega}$ is a vector, even though a finite rotation is not. The relation between the angular velocity vector and the motion of the parts of a rigid body,

$$\vec{\omega} \times \left(\vec{r}_j - \vec{r}_0\right) \tag{30}$$

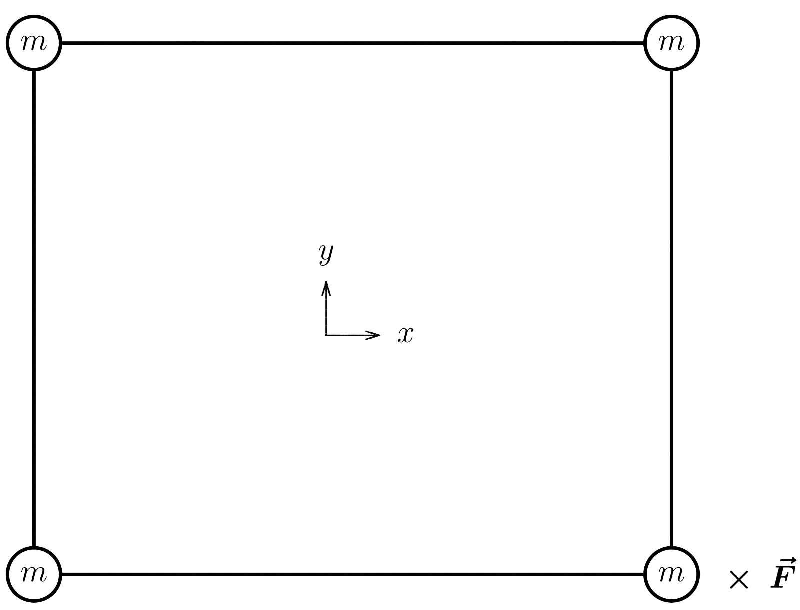
will be the crucial simple fact to hang onto as we explore the complicated world of rigid body rotations. This depends on not only the direction of the axis, but exactly where the axis is, which can be specified by specifying any point $\vec{r}_0$ on the axis. Any other point on the axis gives the same result, because, any other point has the form

$$\vec{r}_a = \vec{r}_0 + a\hat{n} \tag{31}$$

The extra term proportional to $\hat{n}$ doesn't affect (30) because $\vec{\omega}$ and $\hat{n}$ are in the same direction so their cross product vanishes. You should remember that if you are going to add angular velocities, they must be defined with respect to the same reference point $\vec{r}_0$. Otherwise, things get messy.

Impulse and rigid bodies

So to follow up the notion that thinking about impulsive torques is easier to understand, we are going to spend some time dealing with the following question. Suppose that we are out in space, and we come upon a light rigid rectangular frame with masses at the corners, as shown:



(32)

It is floating at rest, say in the x-y plane, when we hit one of the masses with a hammer, applying a large force in the $-z$ direction for a very short time, as shown. This produces an impulsive force and torque on the rigid frame. The question is, immediately after the hammer blow, what is the velocity of each of the masses? We will not actually solve this problem until next time, but today, we will start to explore a crucial component of the answer — we will discuss the moment of inertia tensor and the relation between the angular momentum $\vec{L}$ and the angular velocity $\vec{\omega}$.

But first you might ask — Why have I formulated the problem in this peculiar way? Why not just ask for the trajectory of each of the masses for all time after the hammer blow? The answer is that these trajectories are MUCH harder to find and to understand than the velocities I have asked about. I will begin by explaining why this is so. You may guess that the answer has something to do with impulse — the fact that we have applied the force in a very short time so that the frame does not have a chance to move while the force is being applied. That is correct, but it is only part of the difference. Complications arise because the direction of the angular momentum after the hammer blow does not coincide with the instantaneous axis of rotation of the body. When we calculate $\vec{L}$ in terms of $\vec{\omega}$ and the parameters m_j and $\vec{r}_j$ that describe the rigid body, it just turns out that except in very special circumstances, $\vec{L}$ and $\vec{\omega}$ are not in the same direction. This fact will cause us lots of grief when we try to calculate the actual trajectories, and I want to postpone the worst of it. In fact, in this course, we will not actually ever solve for the full trajectories in this case, although we will do so in some interesting and very non-trivial examples. But just finding the velocities of the masses right after the hammer strike is not so bad. We do this by studying carefully the vectors $\vec{p}$ (linear momentum), $\vec{L}$ (angular momentum), and $\vec{\omega}$ (angular velocity) and understanding the relationships between them in detail. We will find a peculiar relationship depending on a complicated object called the moment of inertia tensor. We will explain this as cleverly as we can, but this is one of those cases in which no amount of cleverness can make it look really simple. There are times when you just have to be very careful and let the mathematics carry you along. I hope that going over it in several different ways will help you get a feeling for this difficult subject.

The moment of inertia tensor

Let's go back to our expression for the angular momentum of a rigid body about a point $\vec{R}$,

$$\vec{L} = \sum_j (\vec{r}_j - \vec{R}\,) \times \vec{p}_j = \sum_j m_j\,(\vec{r}_j - \vec{R}\,) \times (\dot{\vec{r}}_j - \dot{\vec{R}}\,) \tag{33}$$

where the sum runs over the various massive parts of the system, labeled by the index j. We will most often be interested in the case where $\vec{R}$ is the center of mass, but that is not necessary. All we are going to assume in (33) is that the motion is entirely due to the rotation of the body about the point $\vec{R}$. That is, we are calculating the angular momentum in a frame in which the point $\vec{R}$ is fixed. Now we can calculate this for a rigid body rotating with angular velocity vector $\vec{\omega}$ around $\vec{R}$ by using

$$(\dot{\vec{r}}_j - \dot{\vec{R}}\,) = \vec{\omega} \times (\vec{r}_j - \vec{R}\,) \tag{34}$$

Putting (34) into (33) gives

$$\vec{L} = \sum_j m_j\,(\vec{r}_j - \vec{R}\,) \times \Big(\vec{\omega} \times (\vec{r}_j - \vec{R}\,)\Big) \tag{35}$$

Notice that the components of $\vec{L}$ are just linear combinations of the components of $\vec{\omega}$, but the coefficients are some complicated looking sums.

To get a feel for these scary looking sums, let us begin by looking at a case when there is just a single term.

$$\vec{L} = m_1\,(\vec{r}_1 - \vec{R}\,) \times \Big(\vec{\omega} \times (\vec{r}_1 - \vec{R}\,)\Big) \tag{36}$$

This describes the angular momentum of a single particle of mass m_1 at the point $\vec{r}_1$ rotating about the point $\vec{R}$ with angular velocity $\vec{\omega}$. And let's simplify further by putting $\vec{R} = 0$ (this is just a choice of origin anyway, so it doesn't cost anything). Then (36) becomes

$$\vec{L} = m_1\,\vec{r}_1 \times (\vec{\omega} \times \vec{r}_1) = -m_1\,\vec{r}_1 \times (\vec{r}_1 \times \vec{\omega}) \tag{37}$$

We can understand what (37) is by decomposing $\vec{\omega}$ into pieces perpendicular to and parallel to $\vec{r}_1$,

$$\vec{\omega} = \vec{\omega}_\perp + \vec{\omega}_\parallel \tag{38}$$

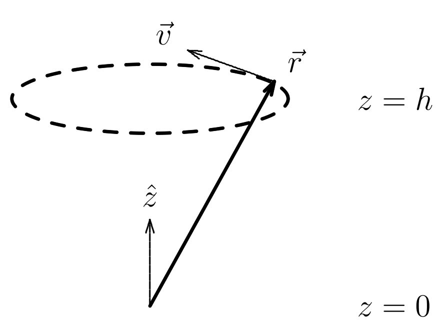
Now the parallel part does not contribute at all because of the cross products, so

$$\vec{L} = -m_1\,\vec{r}_1 \times (\vec{r}_1 \times \vec{\omega}_\perp) \tag{39}$$

But now each of the cross products rotates the vector $\vec{\omega}_\perp$ by 90° in the plane perpendicular to $\vec{r}_1$ and multiplies the magnitude by $|\vec{r}_1|$. Thus if we do this twice we get $-|\vec{r}_1|^2\vec{\omega}_\perp$, so (39) becomes

$$\vec{L} = -m_1\,\vec{r}_1 \times (\vec{r}_1 \times \vec{\omega}_\perp) = m_1\,|\vec{r}_1|^2\vec{\omega}_\perp = m_1\,(\vec{r}_1 \cdot \vec{r}_1)\,\vec{\omega}_\perp \tag{40}$$

Thus the angular momentum is not so complicated in this case. It is just $m_1 r_1^2$ times the perpendicular component of $\vec{\omega}$. However, already in this simple case, $\vec{\omega}$ and $\vec{L}$ are **not parallel!** We saw this in a prs question a few lectures ago.



In the diagram, $\vec{\omega}$ is in the $\hat{z}$ direction, but $\vec{L}$ is perpendicular to $\vec{r}$ in plane formed by $\hat{z}$ and $\vec{r}$. For the motion of a single point mass, while $\vec{\omega}$ determines the velocity of the mass, the velocity does not uniquely determine $\vec{\omega}$. $\vec{\omega}_\perp$ gives the same velocity as $\vec{\omega}$, and it is $\vec{\omega}_\perp$ that matters for $\vec{L}$ because it is perpendicular to $\vec{r}$.

The notation is (40) is not good enough to allow us to generalize this to the case of more than one j. The problem is that $\vec{\omega}_\perp$ depends implicitly on the direction of $\vec{r}_1$. To put this back into (35) we have to make this dependence explicit. To do that, note that the parallel component of $\vec{\omega}$ can be written as

$$\vec{\omega}_\parallel = \hat{r}_1\,(\hat{r}_1 \cdot \vec{\omega}) \tag{41}$$

and thus from (38), we can write

$$\vec{\omega}_\perp = \vec{\omega} - \hat{r}_1\,(\hat{r}_1 \cdot \vec{\omega}) \tag{42}$$

Combining (40) and (42) gives

$$\vec{L} = m_1\,(\vec{r}_1 \cdot \vec{r}_1)\,\vec{\omega}_\perp = m_1\,(\vec{r}_1 \cdot \vec{r}_1)\left(\vec{\omega} - \hat{r}_1\,(\hat{r}_1 \cdot \vec{\omega})\right) = m_1\,(\vec{r}_1 \cdot \vec{r}_1)\,\vec{\omega} - m_1\,\vec{r}_1\,(\vec{r}_1 \cdot \vec{\omega}) \tag{43}$$

Putting the arbitrary reference point $\vec{R}$ back in and factoring out the factor of m_1 gives

$$m_1\left[\left((\vec{r}_1 - \vec{R}\,) \cdot (\vec{r}_1 - \vec{R}\,)\right)\vec{\omega} - (\vec{r}_1 - \vec{R}\,)\left((\vec{r}_1 - \vec{R}\,) \cdot \vec{\omega}\right)\right] \tag{44}$$

Now all the dependence on $\vec{r}_1$ is completely explicit, and we can apply the same procedure to each of the terms in the sum in (35). The result is

$$\vec{L} = \sum_j m_j\left[\left((\vec{r}_j - \vec{R}\,) \cdot (\vec{r}_j - \vec{R}\,)\right)\vec{\omega} - (\vec{r}_j - \vec{R}\,)\left((\vec{r}_j - \vec{R}\,) \cdot \vec{\omega}\right)\right] \tag{45}$$

Equation (45) is the desired relation between $\vec{L}$ and $\vec{\omega}$. However, it is useful to write it in a different form. The first thing to notice is that both sides of (45) are vectors, and that the right hand side is proportional to the vector $\vec{\omega}$, but is not, in general, in the same direction as $\vec{\omega}$. The second thing to notice is that the second term in (45) involves a dot product. We can think of this as the sum that appears in a matrix multiplication. This allows us to write (45) in a slightly more useful form in matrix notation. First define a matrix $\vec{\vec{B}}$ as follows:

$$\vec{\vec{B}} = \sum_j m_j\,(\vec{r}_j - \vec{R}\,)\,(\vec{r}_j - \vec{R}\,) = \begin{pmatrix} B_{xx} & B_{xy} & B_{xz} \\ B_{yx} & B_{yy} & B_{yz} \\ B_{zx} & B_{zy} & B_{zz} \end{pmatrix} \tag{46}$$

$$= \sum_j m_j \begin{pmatrix} [\vec{r}_j - \vec{R}\,]_x\,[\vec{r}_j - \vec{R}\,]_x & [\vec{r}_j - \vec{R}\,]_x\,[\vec{r}_j - \vec{R}\,]_y & [\vec{r}_j - \vec{R}\,]_x\,[\vec{r}_j - \vec{R}\,]_z \\ [\vec{r}_j - \vec{R}\,]_y\,[\vec{r}_j - \vec{R}\,]_x & [\vec{r}_j - \vec{R}\,]_y\,[\vec{r}_j - \vec{R}\,]_y & [\vec{r}_j - \vec{R}\,]_y\,[\vec{r}_j - \vec{R}\,]_z \\ [\vec{r}_j - \vec{R}\,]_z\,[\vec{r}_j - \vec{R}\,]_x & [\vec{r}_j - \vec{R}\,]_z\,[\vec{r}_j - \vec{R}\,]_y & [\vec{r}_j - \vec{R}\,]_z\,[\vec{r}_j - \vec{R}\,]_z \end{pmatrix} \tag{47}$$

We have used a notation of two vector signs to indicated this matrix (Dave, in his book, just uses bold face and you have to figure out from the context whether you are dealing with a vector or a matrix). In general, this is not a good thing to do, but for our purposes in this course, it doesn't do any harm to use a slightly sloppy notation in which we think of our matrices as "bivectors." In fact, it will be kind of helpful. The reason is that the matrices we will have to deal with are symmetric, so we don't have to be careful to distinguish which vector index is which. You can see the symmetry from the definition of $\vec{\vec{B}}$, because

$$B_{xy} = B_{yx}\,, \quad B_{xz} = B_{zx}\,, \quad B_{yz} = B_{zy}\,. \tag{48}$$

Then we can think of the dot product in (45) as matrix multiplication, and just indicate it as a dot product (this is actually the notation that *Mathematica* uses). Then the second term in (45) can be written as

$$-\vec{\vec{B}}\cdot\vec{\omega} \tag{49}$$

The first term in (45) involves the trace of the matrix $\vec{\vec{B}}$

$$\sum_j m_j\,(\vec{r}_j-\vec{R}\,)^2 = \mathrm{Tr}\,\vec{\vec{B}} = B_{xx}+B_{yy}+B_{zz} \tag{50}$$

So we can combine the two terms in (45) as a single matrix equation,

$$\vec{L} = \vec{\vec{I}}\cdot\vec{\omega} \tag{51}$$

where $\vec{\vec{I}}$ is the matrix

$$\begin{pmatrix} B_{xx}+B_{yy}+B_{zz} & 0 & 0 \\ 0 & B_{xx}+B_{yy}+B_{zz} & 0 \\ 0 & 0 & B_{xx}+B_{yy}+B_{zz} \end{pmatrix} - \begin{pmatrix} B_{xx} & B_{xy} & B_{xz} \\ B_{yx} & B_{yy} & B_{yz} \\ B_{zx} & B_{zy} & B_{zz} \end{pmatrix} \tag{52}$$

or

$$\vec{\vec{I}} = \begin{pmatrix} B_{yy}+B_{zz} & -B_{xy} & -B_{xz} \\ -B_{yx} & B_{xx}+B_{zz} & -B_{yz} \\ -B_{zx} & -B_{zy} & B_{xx}+B_{yy} \end{pmatrix} \tag{53}$$

Like the matrix $\vec{\vec{B}}$, the matrix $\vec{\vec{I}}$ is symmetric —

$$I_{xy}=I_{yx}\,, \quad I_{xz}=I_{zx}\,, \quad I_{yz}=I_{zy}\,. \tag{54}$$

Thus $\vec{\vec{I}}$ is completely determined by 6 numbers,

$$I_{xx}\,, \quad I_{yy}\,, \quad I_{zz}\,, \quad I_{xy}\,, \quad I_{xz}\,, \quad \text{and} \quad I_{yz}\,. \tag{55}$$

The object $\vec{\vec{I}}$ is a thing called the "moment of inertia tensor." The term "tensor" refers to a large class of objects in mathematics that are generalizations of vectors. We are not going to get into tensors in general. This particular tensor is fairly simple because it can be thought of as a machine for taking linear combinations of the components of one vector ($\vec{\omega}$) to get another vector ($\vec{L}$). Equations (51) and (53) show that you can also think of $\vec{\vec{I}}$ as a 3×3 matrix, and think of the right hand side of (51) as a matrix product of the matrix $\vec{\vec{I}}$ and the vector $\vec{\omega}$. Because (51) is an equation for a vector, it has a meaning independent of the particular coordinate system in which we describe the vectors. But like a vector, the tensor $\vec{\vec{I}}$ will have different components in different coordinate systems. What makes the moment of inertia tensor a tensor is that when we go from one coordinate system to another, the components of $\vec{\vec{I}}$ transform in the just the right way so that (51) is still a correct vector equation. That is the meaning of the two vector signs on $\vec{\vec{I}}$. This is implicit in (51), in which one of the vector signs on $\vec{\vec{I}}$ is associated with the vector nature of $\vec{L}$ on the left hand side, while the other vector index is combined by the dot product with the vector index of $\vec{\omega}$ so that the right hand side behaves like a vector. This will be very important in our understanding of rigid bodies.

The general point about tensors is that one can actually go backwards through this argument and argue that $\vec{\vec{I}}$ (because it takes a vector into another vector) behaves under rotations like a product of two vectors. That doesn't mean that it is a product of two vectors, just that it rotates the way a product of two vectors rotates when we go from one coordinate system to another.

In term of $\vec{\vec{I}}$, we can quickly derive for rotation about a fixed axis, $\hat{n}$, the component of angular momentum in the $\hat{n}$ direction. In general, the $\hat{n}$ component of $\vec{L}$ is

$$\hat{n} \cdot \vec{L} = \hat{n} \cdot \vec{\vec{I}} \cdot \vec{\omega} \tag{56}$$

but for a body that is rotating around the $\hat{n}$ axis, $\vec{\omega} = \omega\,\hat{n}$ and so

$$\hat{n} \cdot \vec{L} = \hat{n} \cdot \vec{\vec{I}} \cdot \vec{\omega} = \hat{n} \cdot \vec{\vec{I}} \cdot \omega\hat{n} = \omega\,\hat{n} \cdot \vec{\vec{I}} \cdot \hat{n} \tag{57}$$

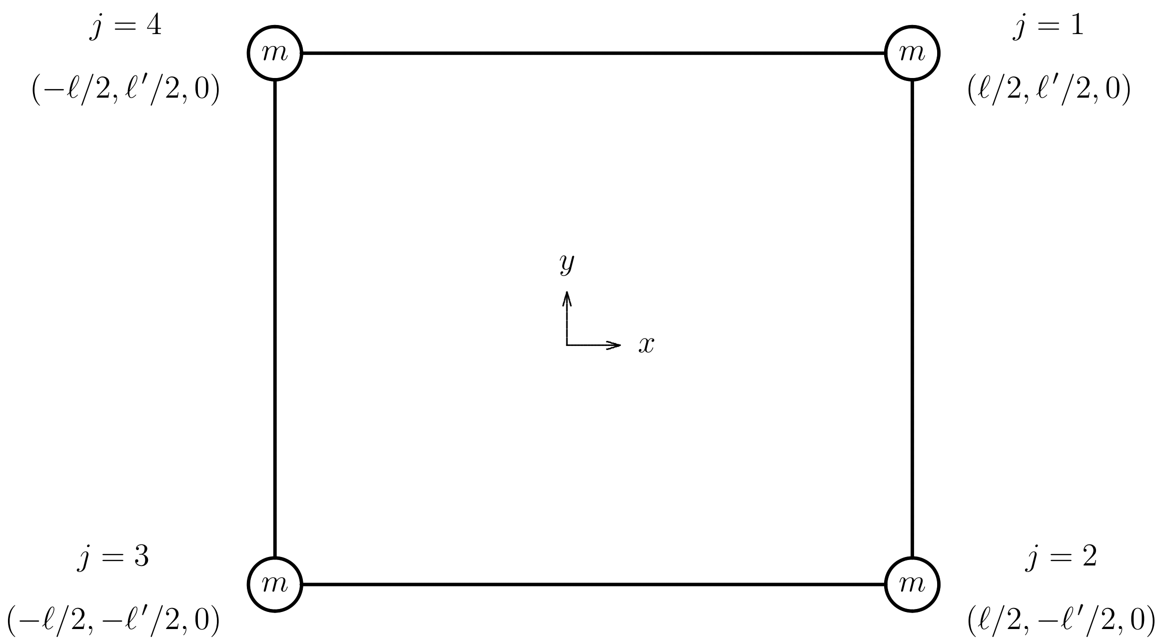
But we showed last week that ω times the moment of inertia about the axis $\hat{n}$, $I_{\hat{n}}$, is ω times the component in the $\hat{n}$ direction of angular momentum associated with angular velocity $\omega\,\hat{n}$. Thus (57) implies that

$$I_{\hat{n}} = \hat{n} \cdot \vec{\vec{I}} \cdot \hat{n} \tag{58}$$

This means that our tensor $\vec{\vec{I}}$ contains all the information about the moment of inertia about any possible axis!

Let's see what $\vec{\vec{I}}$ looks like for the example system if we take the reference point $\vec{R}$ to be the center of mass. If the rectangular frame has sides of length ℓ and ℓ', and we center the object in the

$z = 0$ plane, the positions of the mass are as shown below:



(59)

The center of mass (the vector $\vec{R}$) is at the origin, by symmetry or explicit calculation

$$\vec{R} = \frac{1}{M} \sum_j m_j \, \vec{r}_j = (0, 0, 0) \tag{60}$$

Then we have

$$m_j = m\,, \quad \vec{r}_j - \vec{R} = (\pm\ell/2, \pm\ell'/2, 0) \tag{61}$$

where the $\pm$ signs in (61) are independent, running over all four possibilities as j runs from 1 to 4.

Now we can compute the various components of the moment of inertia tensor: For example

$$\begin{aligned}
\vec{\vec{I}}_{xx} &= \sum_j m_j \left((\vec{r}_j - \vec{R})^2 - (r_j - R)_x^2 \right) \\
&= \sum_j m_j \left((r_j - R)_y^2 + (r_j - R)_z^2 \right) = m \, \ell'^{\,2} \\
\vec{\vec{I}}_{yy} &= \sum_j m_j \left((r_j - R)_x^2 + (r_j - R)_z^2 \right) = m \, \ell^2 \\
\vec{\vec{I}}_{xy} = \vec{\vec{I}}_{yx} &= -\sum_j m_j \left((r_j - R)_x (r_j - R)_y \right) = 0 \\
\vec{\vec{I}}_{zz} = \vec{\vec{I}}_{xx} + \vec{\vec{I}}_{yy} & \qquad \vec{\vec{I}}_{xz} = \vec{\vec{I}}_{yz} = \vec{\vec{I}}_{zx} = \vec{\vec{I}}_{zy} = 0
\end{aligned} \tag{62}$$

In matrix form, this looks like

$$\vec{\vec{I}} = \begin{pmatrix} m\ell'^2 & 0 & 0 \\ 0 & m\ell^2 & 0 \\ 0 & 0 & m(\ell^2 + \ell'^2) \end{pmatrix} \tag{63}$$

Notice that the vanishing of $\vec{\vec{I}}_{xy}$ is slightly of non-trivial. It follows from the form of (61), but requires a cancelation between different terms in the sum. This happened only because we were clever (or at least not stupid) in our choice of coordinate system.

An impulsive demo

An impulsive force is one that is applied for a very short time. To get a large change in momentum in a very short time requires a very large impulsive force. Applying a very large force is a good way of breaking things. Here is an example.

lecture 19

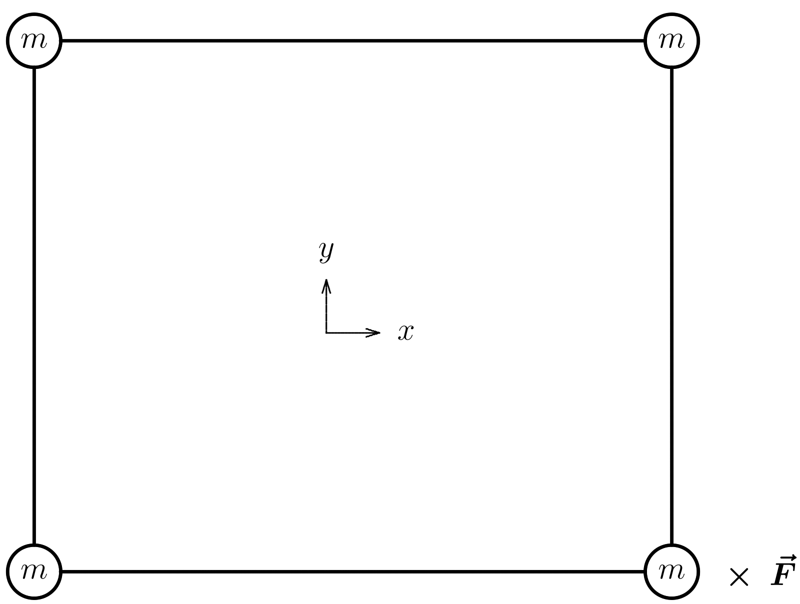
Topics:

Where are we now?
The moment of inertia tensor $\vec{\vec{I}}$
$\vec{L}$ and $\vec{\omega}$
The body frame

Where are we now?

We are poised to take the next big step into understanding the motion of rigid bodies. By putting together the ideas of impulsive torque and the angular velocity vector, we can solve some more interesting problems. Last time, we introduced the moment of inertia tensor. This time we will briefly review this and go on to apply it in interesting ways.

Last time, we introduced the following problem. Suppose that we are out in space, and we come upon a light rigid rectangular frame with masses at the corners, as shown:



(1)

It is floating at rest, say in the x-y plane, when we hit one of the masses with a hammer, applying a large force in the $-z$ direction for a very short time, as shown. This produces an impulsive force and torque on the rigid frame. The question is, immediately after the hammer blow, what is the velocity of each of the masses? We began this analysis last time by defining the moment of inertia tensor and the relation between the angular momentum $\vec{L}$ and the angular velocity $\vec{\omega}$.

The moment of inertia tensor

Let's begin by reviewing what we said last time about the moment of inertia tensor. It may be comforting to note that the scary definition we arrived at last time, and will look at again today is not the way we will usually calculate $\vec{\vec{I}}$. At the end of today's lecture, we will begin to discuss easier ways. But it is important to understand where it comes from so we will review it and apply it to our floating frame problem.

Let's go back to our expression for the angular momentum of a rigid body about a point $\vec{R}$,

$$\vec{L} = \sum_j (\vec{r}_j - \vec{R}\,) \times \vec{p}_j = \sum_j m_j\,(\vec{r}_j - \vec{R}\,) \times (\dot{\vec{r}}_j - \dot{\vec{R}}\,) \tag{2}$$

We showed last time that we can rewrite this as

$$\vec{L} = \sum_j m_j \left[\Big((\vec{r}_j - \vec{R}\,) \cdot (\vec{r}_j - \vec{R}\,) \Big)\, \vec{\omega} - (\vec{r}_j - \vec{R}\,) \Big((\vec{r}_j - \vec{R}\,) \cdot \vec{\omega} \Big) \right] \tag{3}$$

Equation (3) is the desired relation between $\vec{L}$ and $\vec{\omega}$. However, it is useful to write it in a different form. The first thing to notice is that both sides of (3) are vectors, and that the right hand side is proportional to the vector $\vec{\omega}$, but is not, in general, in the same direction as $\vec{\omega}$. The second thing to notice is that the second term in (3) involves a dot product. We can think of this as the sum that appears in a matrix multiplication. This allows us to write (3) in a slightly more useful form in matrix notation.

First define a matrix $\vec{\vec{B}}$ as follows:

$$\vec{\vec{B}} = \sum_j m_j\,(\vec{r}_j - \vec{R}\,)\,(\vec{r}_j - \vec{R}\,) = \begin{pmatrix} B_{xx} & B_{xy} & B_{xz} \\ B_{yx} & B_{yy} & B_{yz} \\ B_{zx} & B_{zy} & B_{zz} \end{pmatrix} \tag{4}$$

$$= \sum_j m_j \begin{pmatrix} [\vec{r}_j - \vec{R}\,]_x\,[\vec{r}_j - \vec{R}\,]_x & [\vec{r}_j - \vec{R}\,]_x\,[\vec{r}_j - \vec{R}\,]_y & [\vec{r}_j - \vec{R}\,]_x\,[\vec{r}_j - \vec{R}\,]_z \\ [\vec{r}_j - \vec{R}\,]_y\,[\vec{r}_j - \vec{R}\,]_x & [\vec{r}_j - \vec{R}\,]_y\,[\vec{r}_j - \vec{R}\,]_y & [\vec{r}_j - \vec{R}\,]_y\,[\vec{r}_j - \vec{R}\,]_z \\ [\vec{r}_j - \vec{R}\,]_z\,[\vec{r}_j - \vec{R}\,]_x & [\vec{r}_j - \vec{R}\,]_z\,[\vec{r}_j - \vec{R}\,]_y & [\vec{r}_j - \vec{R}\,]_z\,[\vec{r}_j - \vec{R}\,]_z \end{pmatrix} \tag{5}$$

We have used a notation of two vector signs to indicated this matrix (Dave, in his book, just uses bold face and you have to figure out from the context whether you are dealing with a vector or a matrix). In general, this is not a good thing to do, but for our purposes in this course, it doesn't do any harm to use a slightly sloppy notation in which we think of our matrices as "bivectors." In fact, it will be kind of helpful. The reason is that the matrices we will have to deal with are symmetric, so we don't have to be careful to distinguish which vector index is which. You can see the symmetry from the definition of $\vec{\vec{B}}$, because

$$B_{xy} = B_{yx}\,, \quad B_{xz} = B_{zx}\,, \quad B_{yz} = B_{zy}\,. \tag{6}$$

So we can combine the two terms in (3) as a single matrix equation,

$$\vec{L} = \vec{\vec{I}} \cdot \vec{\omega} \tag{7}$$

where $\vec{\vec{I}}$ is the matrix

$$\begin{pmatrix} B_{xx}+B_{yy}+B_{zz} & 0 & 0 \\ 0 & B_{xx}+B_{yy}+B_{zz} & 0 \\ 0 & 0 & B_{xx}+B_{yy}+B_{zz} \end{pmatrix} - \begin{pmatrix} B_{xx} & B_{xy} & B_{xz} \\ B_{yx} & B_{yy} & B_{yz} \\ B_{zx} & B_{zy} & B_{zz} \end{pmatrix} \tag{8}$$

or

$$\vec{\vec{I}} = \begin{pmatrix} B_{yy}+B_{zz} & -B_{xy} & -B_{xz} \\ -B_{yx} & B_{xx}+B_{zz} & -B_{yz} \\ -B_{zx} & -B_{zy} & B_{xx}+B_{yy} \end{pmatrix} \tag{9}$$

Like the matrix $\vec{\vec{B}}$, the matrix $\vec{\vec{I}}$ is symmetric —

$$I_{xy} = I_{yx} \,, \quad I_{xz} = I_{zx} \,, \quad I_{yz} = I_{zy} \,. \tag{10}$$

Thus $\vec{\vec{I}}$ is completely determined by 6 numbers,

$$I_{xx} \,, \quad I_{yy} \,, \quad I_{zz} \,, \quad I_{xy} \,, \quad I_{xz} \,, \quad \text{and} \quad I_{yz} \,. \tag{11}$$

The object $\vec{\vec{I}}$ is a thing called the "moment of inertia tensor." The term "tensor" refers to a large class of objects in mathematics that are generalizations of vectors. I was not planning on talking about tensors in general, but there were so many questions about the concept on the QA, that I felt I had to say something about it. So here is a definition of "tensor" that makes sense to a physicist (mathematicians would say things much more abstractly). **A tensor is anything that can be written as a linear combination (that is just a sum with coefficients) products of vectors.** We have talked a lot about what makes a vector a vector, which is the way it transforms under a symmetry transformation. And we know that there are different kinds of vectors depending on what symmetry we are talking about. The 3-vectors you are familiar with from high school transform under rotations in 3-dimensional space, but we have also now seen 4-vectors that transform under Lorentz transformations. And there are other things are useful as well. The moment of inertia tensor is a sum of products of ordinary 3-vectors. There are different kinds of tensors depending on how many vectors you multiply together. The number of vectors in the product is called the "rank" of the tensor. A vector is a rank 1 tensor. Just like with vectors, you can add tensors of the same rank together and they are still tensors of the same rank, and you can write tensor equations in which both sides are tensors of the same rank. And indeed this is a good thing to do, because then if the equation is true in one coordinate system, it is automatically correct in all coordinate systems.

You can see how this works explicitly in (4) and (5) where you see that each component of $\vec{\vec{B}}$ (which is a rank 2 tensor) is a linear combination of products of components of the vectors $\vec{r}_j - \vec{R}$.

The reason that I didn't want to spend time on this is that it is a little more complicate to explain why the first term in (8),

$$\vec{\vec{A}} = \begin{pmatrix} B_{xx}+B_{yy}+B_{zz} & 0 & 0 \\ 0 & B_{xx}+B_{yy}+B_{zz} & 0 \\ 0 & 0 & B_{xx}+B_{yy}+B_{zz} \end{pmatrix} \tag{12}$$

is a tensor. This doesn't look like a linear combination of products of two vectors, but actually we can write it as

$$\vec{\vec{A}} = (B_{xx}+B_{yy}+B_{zz})(\hat{x}\,\hat{x}+\hat{y}\,\hat{y}+\hat{z}\,\hat{z}) \tag{13}$$

This is special and important feature of the identity matrix. So in fact, both terms in (8) are linear combinations of products of two vectors, and so $\vec{\vec{I}}$ is a rank 2 tensor.

Rank 2 tensors, like the moment of inertia tensor, are special because they can be written as matrices and thought of as a machine for taking linear combinations of the components of one vector ($\vec{\omega}$) to get another vector ($\vec{L}$). "Matrix" is just a generic term for something with rows and columns. The particular matrices that we care about are rank 2 tensors. Equations (7) and (9) show that you can also think of $\vec{\vec{I}}$ as a 3×3 matrix, and think of the right hand side of (7) as a matrix product of the matrix $\vec{\vec{I}}$ and the vector $\vec{\omega}$. Because (7) is an equation for a vector, it has a meaning independent of the particular coordinate system in which we describe the vectors. But like a vector, the tensor $\vec{\vec{I}}$ will have different components in different coordinate systems. What makes the moment of inertia tensor a tensor is that when we go from one coordinate system to another, the components of $\vec{\vec{I}}$ transform in the just the right way so that (7) is still a correct vector equation. That is the meaning of the two vector signs on $\vec{\vec{I}}$. This is implicit in (7), in which one of the vector signs on $\vec{\vec{I}}$ is associated with the vector nature of $\vec{L}$ on the left hand side, while the other vector index is combined by the dot product with the vector index of $\vec{\omega}$ so that the right hand side behaves like a vector. Incidentally, for the mathematically inclined, note that taking a dot product in this way is general way of reducing the rank of a tensor. If we just multiplied the 9 components of $\vec{\vec{I}}$ with the 3 components of $\vec{\omega}$ in all possible ways, we would get 27 components of a rank 3 tensor

$$\vec{\vec{\vec{C}}} = \vec{\vec{I}}\,\vec{\omega} \tag{14}$$

But by taking the dot product, we can get rid of two vectors and make a rank 1 tensor - that is a vector. This will be very important in our understanding of rigid bodies.

In term of $\vec{\vec{I}}$, we can quickly derive for rotation about a fixed axis, $\hat{n}$, the component of angular momentum in the $\hat{n}$ direction. In general, the $\hat{n}$ component of $\vec{L}$ is

$$\hat{n} \cdot \vec{L} = \hat{n} \cdot \vec{\vec{I}} \cdot \vec{\omega} \tag{15}$$

but for a body that is rotating around the $\hat{n}$ axis, $\vec{\omega} = \omega\,\hat{n}$ and so

$$\hat{n} \cdot \vec{L} = \hat{n} \cdot \vec{\vec{I}} \cdot \vec{\omega} = \hat{n} \cdot \vec{\vec{I}} \cdot \omega\hat{n} = \omega\,\hat{n} \cdot \vec{\vec{I}} \cdot \hat{n} \tag{16}$$

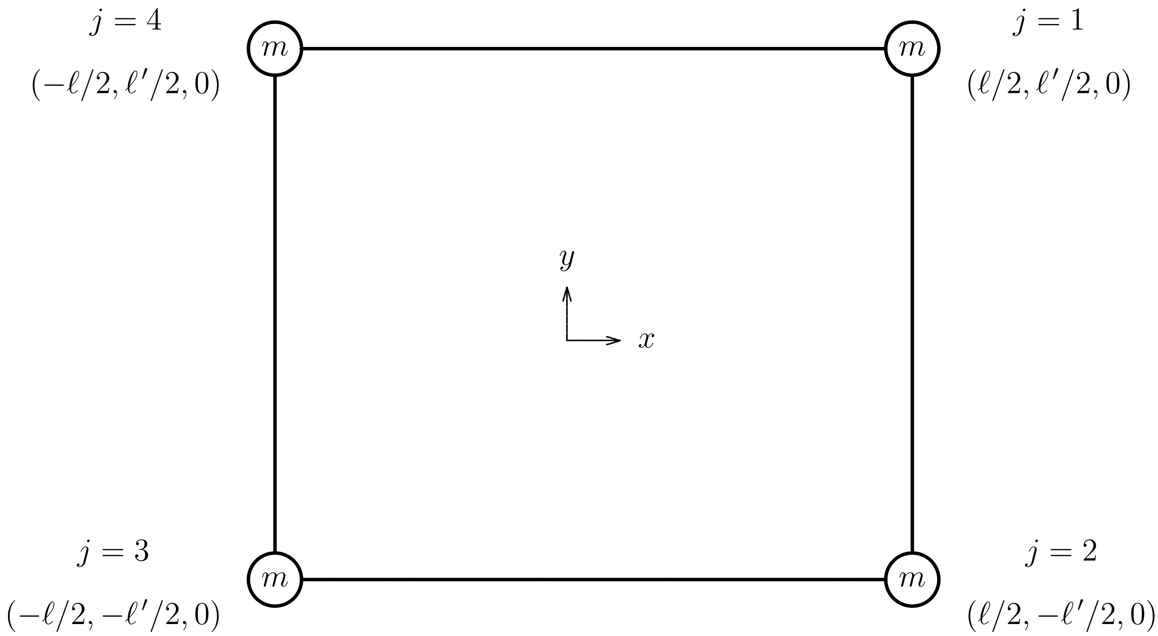
But we showed last week that ω times the moment of inertia about the axis $\hat{n}$, $I_{\hat{n}}$, is ω times the component in the $\hat{n}$ direction of angular momentum associated with angular velocity $\omega\,\hat{n}$. Thus (16) implies that

$$I_{\hat{n}} = \hat{n} \cdot \vec{\vec{I}} \cdot \hat{n} \tag{17}$$

This means that our tensor $\vec{\vec{I}}$ contains all the information about the moment of inertia about any possible axis!

Let's see what $\vec{\vec{I}}$ looks like for the example system if we take the reference point $\vec{R}$ to be the center of mass. If the rectangular frame has sides of length ℓ and ℓ', and we center the object in the

$z = 0$ plane, the positions of the mass are as shown below:



(18)

The center of mass (the vector $\vec{R}$) is at the origin, by symmetry or explicit calculation

$$\vec{R} = \frac{1}{M} \sum_j m_j \, \vec{r}_j = (0, 0, 0) \tag{19}$$

Then we have

$$m_j = m \,, \quad \vec{r}_j - \vec{R} = (\pm\ell/2, \pm\ell'/2, 0) \tag{20}$$

where the $\pm$ signs in (20) are independent, running over all four possibilities as j runs from 1 to 4.

Now we can compute the various components of the moment of inertia tensor: For example

$$\begin{aligned}
\vec{\vec{I}}_{xx} &= \sum_j m_j \left((\vec{r}_j - \vec{R}\,)^2 - (r_j - R)_x^2 \right) \\
&= \sum_j m_j \left((r_j - R)_y^2 + (r_j - R)_z^2 \right) = m \, \ell'^{\,2} \\
\vec{\vec{I}}_{yy} &= \sum_j m_j \left((r_j - R)_x^2 + (r_j - R)_z^2 \right) = m \, \ell^{\,2} \\
\vec{\vec{I}}_{xy} = \vec{\vec{I}}_{yx} &= -\sum_j m_j \left((r_j - R)_x (r_j - R)_y \right) = 0 \\
\vec{\vec{I}}_{zz} = \vec{\vec{I}}_{xx} + \vec{\vec{I}}_{yy} & \qquad \vec{\vec{I}}_{xz} = \vec{\vec{I}}_{yz} = \vec{\vec{I}}_{zx} = \vec{\vec{I}}_{zy} = 0
\end{aligned} \tag{21}$$

In matrix form, this looks like

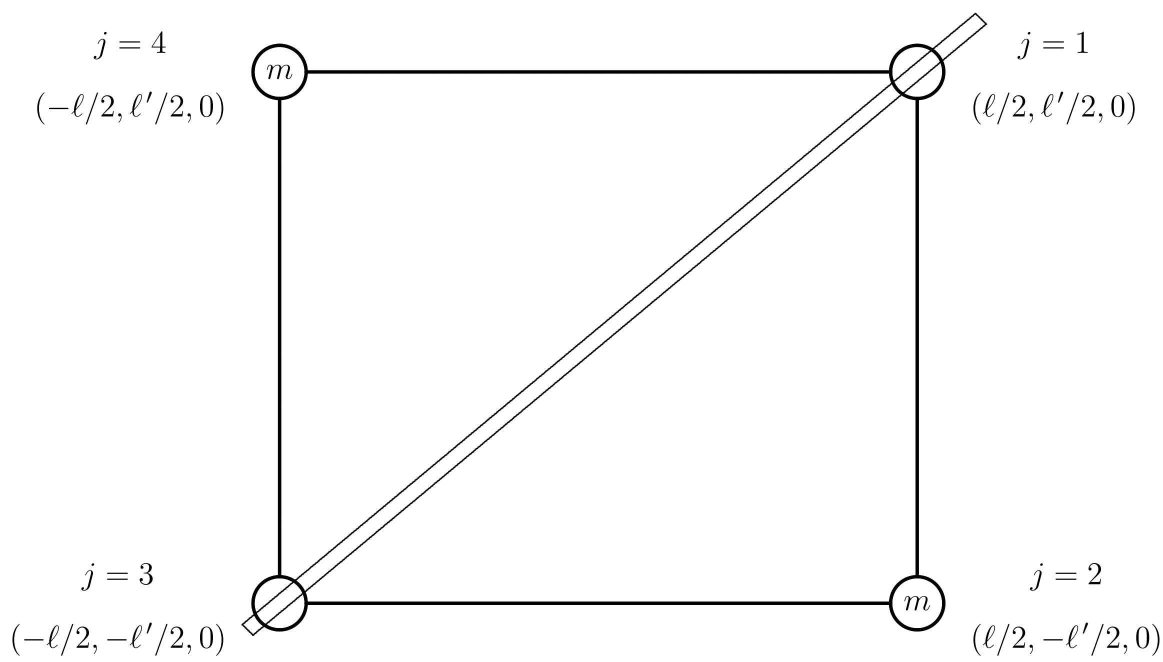
$$\vec{\vec{I}} = \begin{pmatrix} m\ell'^2 & 0 & 0 \\ 0 & m\ell^2 & 0 \\ 0 & 0 & m(\ell^2+\ell'^2) \end{pmatrix} \tag{22}$$

Notice that the vanishing of $\vec{\vec{I}}_{xy}$ is slightly of non-trivial. It follows from the form of (20), but requires a cancelation between different terms in the sum. This happened only because we were clever (or at least not stupid) in our choice of coordinate system.

$\vec{L}$ and $\vec{\omega}$

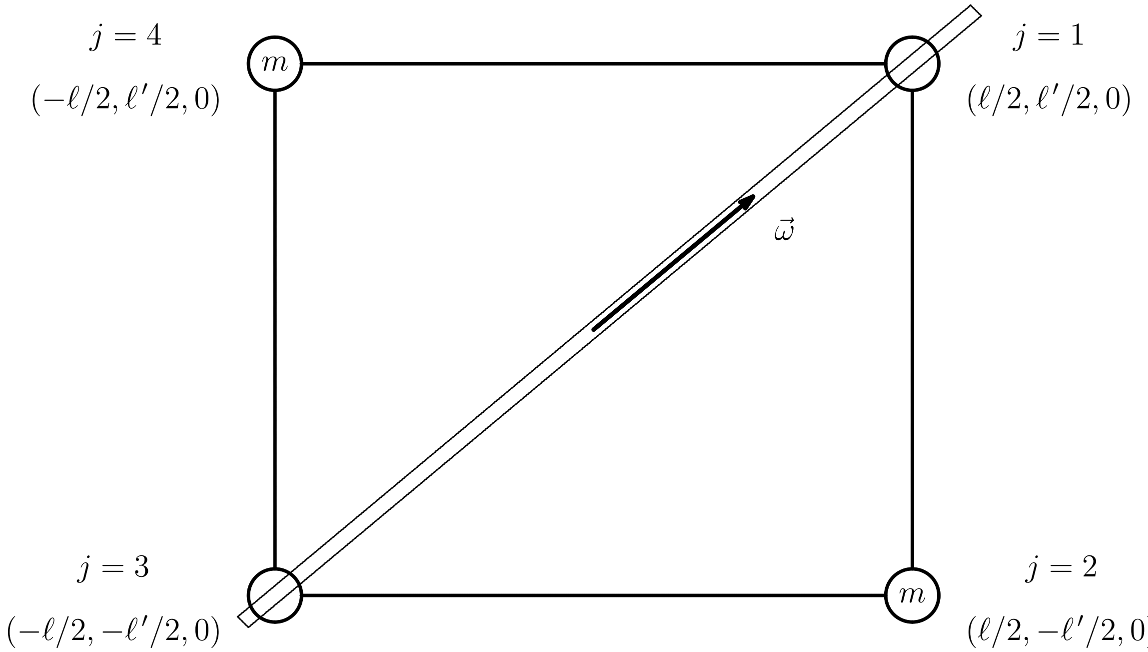
The important thing about the moment of inertia tensor is equation (7), which describes the connection between the angular momentum vector and the angular velocity vector. Because $\vec{\vec{I}}$ is a tensor, and not simply a number, (7) implies that these two vectors are generally not in the same direction. There is a very nontrivial relation between them that depends on the form of $\vec{\vec{I}}$.

One example of the non-trivial connection between $\vec{L}$ and $\vec{\omega}$ is the fact that if one of these vectors is constant, the other is generally changing with time. For example, imagine that we take the rectangular frame we have been discussing and run a frictionless rod through the centers of masses 1 and 3, as shown below



(23)

If we now set the frame rotating about the fixed rod by pushing mass 2 into the plane, the system will have an angular velocity vector along the rod, as shown below:



(24)

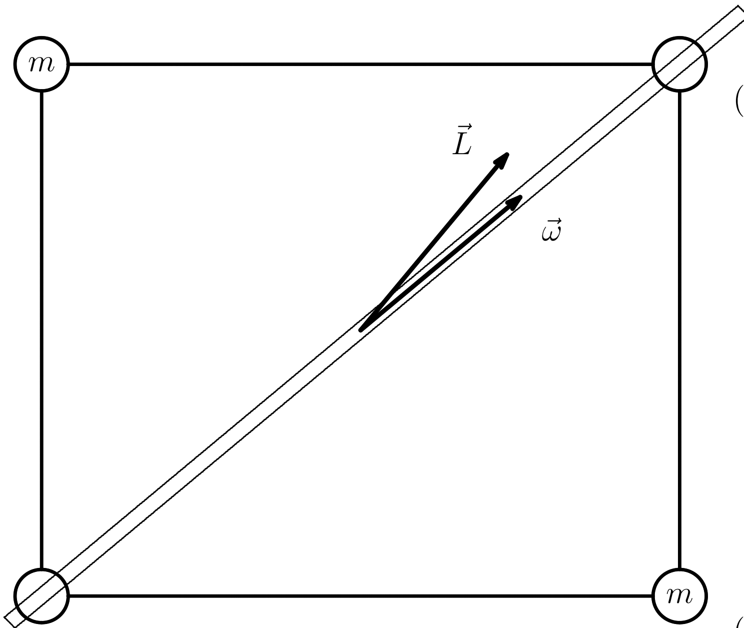
Here $\vec{\omega}$ is in the direction $(\ell, \ell', 0)$, and in the figure, $\ell > \ell'$. But to find the x and y components of $\vec{L}$, we multiply $\vec{\omega}$ by $\vec{\vec{I}}$,

$$\vec{L} = \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} m\ell'^2 & 0 & 0 \\ 0 & m\ell^2 & 0 \\ 0 & 0 & m(\ell^2 + \ell'^2) \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} m\ell'^2\omega_x \\ m\ell^2\omega_y \\ m(\ell^2 + \ell'^2)\omega_z \end{pmatrix} \tag{25}$$

In this case, $\omega_z = 0$, so this has the effect of multiplying the x and y components of $\vec{\omega}$ by different constants,

$$L_x = \vec{\vec{I}}_{xx}\omega_x = m\ell'^2\omega_x\,, \qquad L_y = \vec{\vec{I}}_{yy}\omega_x = m\ell^2\omega_y\,, \tag{26}$$

and this implies that $\vec{L}$ is in a different direction, proportional to $(\ell', \ell, 0)$, as shown below:

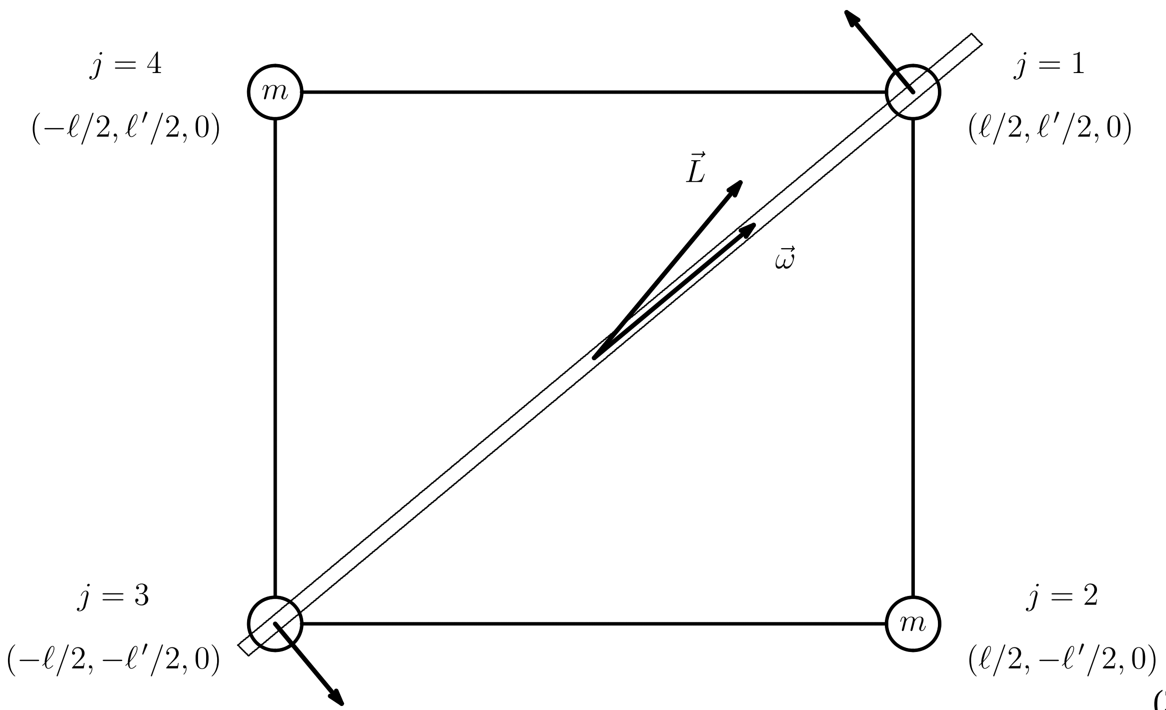


(27)

Now as the frame rotates about the frictionless rod, the angular momentum vector rotates with it! The component of $\vec{L}$ in the direction of the frictionless rod, $L_{\parallel}$, does not change. Because there is no friction, there can be no torque in the direction of the rod, and thus the rate of change of this component of $\vec{L}$ vanishes. That in turn means that the magnitude of $\vec{\omega}$ remains constant as well. But the rod can and does produce forces on the frame perpendicular to the rod. These forces, shown below, produce the torque that causes the perpendicular component of the angular momentum vector, $\vec{L}_{\perp}$, to change. This component executes uniform circular motion with angular frequency $\omega = |\vec{\omega}|$ in the plane perpendicular to $\vec{\omega}$. Thus the magnitude of the torque is

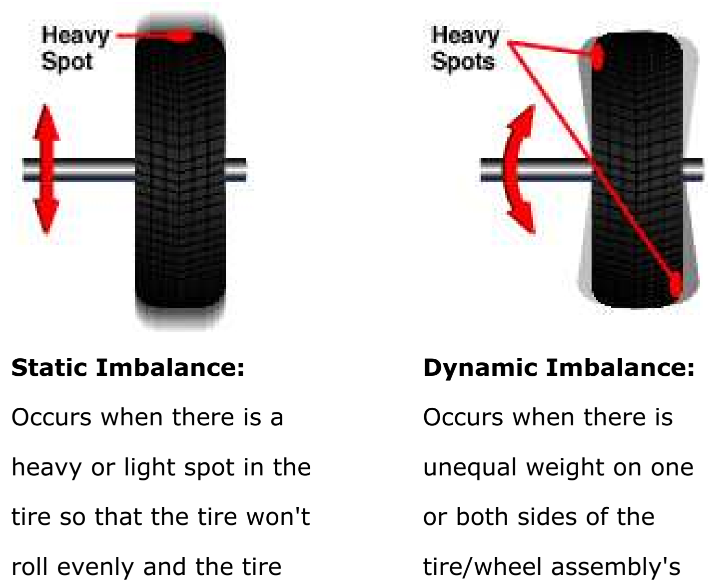
$$|\vec{\omega} \times \vec{L}| = \omega |L_{\perp}| \tag{28}$$

This business of angular momentum and angular velocity being in different directions may not make obvious sense to you. But you should be able to feel this torque in your bones. You should be able to see that the accelerations required to keep this system in uniform circular motion are in different planes, and that this must produce a torque. Since torque is rate of change of angular momentum, you can go backwards and conclude that the angular momentum does not point along the angular velocity. I hope this will help you develop intuition for the meaning of the moment of inertia tensor.



(29)

A practical example of the torque that result from angular momentum that is not in the direction of the angular velocity is the process of balancing tires for your car. Here is a figure from the web (at http://www.discounttire.com/dtc/brochure/info/tireBalance.jsp):



The reason that a mechanic spins your tire during the balancing process is check for dynamic imbalance, which results from mass that is balanced around the center of mass but is not symmetrical with respect to the axis. When a tire with this imbalance spins on the axle, the angular momentum is not in the direction of the axle and the perpendicular component of $\vec{L}$ undergoes uniform circular motion with the tire. Thus there is a non-zero torque, $\vec{\omega} \times \vec{L}$ which causes the tire to shimmy.

Now back in the original problem of hitting the frame while it is floating in space, I hope that you can now begin to see why in general it is hard to find the trajectories of our masses after the system starts to rotate. In this case, after the hit, it is the angular momentum vector that is fixed, rather than the angular velocity vector, but the relation between them is still changing in time.

Let me say this once more in more generality. If a rigid body is rotating freely, the angular momentum vector $\vec{L}$ is fixed in space (because this is an inertial frame - often called the **space frame**),, but the $\vec{r}_j$s that determine the position of the masses in the rigid body are determined by the orientation of the body, so they are all rotating about the center of mass and constantly changing in some way that depends on the angular velocity vector ω. We can calculate ω from $\vec{L}$ if we know the moment of inertia tensor, $\vec{\vec{I}}$. But to find $\vec{\vec{I}}$, we need to know the $\vec{r}_j$s that depend on the orientation of the body. Therefore the form of the moment of inertia tensor is changing in time. It depends on the orientation of the body, which itself is constantly changing. But this means that $\vec{\omega}$ is constantly changing. So this is an incredible mess, and it seems circular. What are we going to do to extricate ourselves?

The body frame

The way out of this dilemma is to think about a frame in which the moment of inertia tensor remains simple. The moment of inertia tensor is constant in a coordinate system that is fixed on the body, and rotates with it. The moment of inertia must be constant in this frame because $\vec{\vec{I}}$ is determined by the masses and position vectors that describe their positions within the body, and if the body is at rest, these position vectors are all constant. This is called the **body frame**. The body frame is certainly not necessarily an inertial frame. In a week or so, we will see how this modifies Newton's laws in such a frame. But even though it is not an inertial frame, it will be useful to us in our analysis of rigid body rotations.

In the body frame, there is a particularly simple choice of coordinate system that makes the moment of inertial tensor look simple. For every body, there are three perpendicular directions, described by unit vectors $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$, which have the nice property that for a rotation about an axis in one of these directions, the angular momentum is in the same direction as the axis of rotation. This is a general result from linear algebra.[1] The axes in the directions $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$ are called the principal axes of the body. Then we say that the body has a moment of inertia I_1 about the axis $\hat{e}_1$, a moment of inertia I_2 about the axis $\hat{e}_2$, and a moment of inertia I_3 about the axis $\hat{e}_3$. We can write the tensor $\vec{\vec{I}}$ in any coordinate system as a sum over the three principal axes,

$$\vec{\vec{I}} = I_1\, \hat{e}_1\, \hat{e}_1 + I_2\, \hat{e}_2\, \hat{e}_2 + I_3\, \hat{e}_3\, \hat{e}_3 \tag{30}$$

[1] Formally, this is the statement that we can always choose a coordinate system in which a given symmetric matrix is diagonal.

If we choose a coordinate system in which $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$ are the basis vectors, then $\vec{\vec{I}}$ is diagonal.

$$\vec{\vec{I}} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{31}$$

The moments I_j around the principal axes are called the principal moments. The directions of the three principal axes completely specify the orientation of the body. The principal axes in our example above are the x, y and z axes.

In matrix language, the principal axes have the following properties:

$$\hat{e}_1 \cdot \hat{e}_1 = \hat{e}_2 \cdot \hat{e}_2 = \hat{e}_3 \cdot \hat{e}_3 = 1 \,, \quad \hat{e}_1 \cdot \hat{e}_2 = \hat{e}_1 \cdot \hat{e}_3 = \hat{e}_2 \cdot \hat{e}_3 = 0 \,, \tag{32}$$

and

$$\vec{\vec{I}} \cdot \hat{e}_1 = I_1 \, \hat{e}_1 \,, \quad \vec{\vec{I}} \cdot \hat{e}_2 = I_2 \, \hat{e}_2 \,, \quad \vec{\vec{I}} \cdot \hat{e}_3 = I_3 \, \hat{e}_3 \,. \tag{33}$$

This is the mathematical statement of the physical properties that I just described. In linear algebra lingo, the vectors $\hat{e}_j$ that describe the principal axes are the **eigenvectors** of the matrix $\vec{\vec{I}}$, and the principal moments, I_j, are the corresponding **eigenvalues**.

For rotation about a principal axis, the angular momentum is in the same direction as the angular velocity and the moment of inertia is the principal moment. For example for rotation about the axis $\hat{e}_1$, the angular velocity has the form $\vec{\omega} = \omega \, \hat{e}_1$, so (33) gives

$$\vec{L} = \vec{\vec{I}} \cdot \vec{\omega} = \omega \, \vec{\vec{I}} \cdot \hat{e}_1 = \omega \, I_1 \, \hat{e}_1 = I_1 \, \vec{\omega} \tag{34}$$

so that indeed the angular momentum is just the principal moment times the angular velocity vector. You can also see explicitly from (30)-(33) that the principal moments are the moments of inertia about the principal axes:

$$\hat{e}_1 \cdot \vec{\vec{I}} \cdot \hat{e}_1 = I_1 \,, \quad \hat{e}_2 \cdot \vec{\vec{I}} \cdot \hat{e}_2 = I_2 \,, \quad \hat{e}_3 \cdot \vec{\vec{I}} \cdot \hat{e}_3 = I_3 \,. \tag{35}$$

This gives us an easy way to calculate the principal moments once we know the principal axes. We can use the expression for the moment of inertia around a fixed axis

$$\hat{n} \cdot \vec{\vec{I}} \cdot \hat{n} = I_{\hat{n}} = \sum_j m_j \, |\hat{n} \times (\vec{r}_j - \vec{R}\,)|^2 \tag{36}$$

to write

$$\begin{aligned} I_1 &= \sum_j m_j \, |\hat{e}_1 \times (\vec{r}_j - \vec{R}\,)|^2 \\ I_2 &= \sum_j m_j \, |\hat{e}_2 \times (\vec{r}_j - \vec{R}\,)|^2 \\ I_3 &= \sum_j m_j \, |\hat{e}_3 \times (\vec{r}_j - \vec{R}\,)|^2 \end{aligned} \tag{37}$$

If the three principal moments of the object have different values, then the principal axes are unique (we will see later what happens with principal moments are equal). In this case, the principal axes are a kind of built in coordinate system in the body frame. Giving the directions of the principal axes in space specifies the orientation of the body in space.

Note that in the impulse problem that we started at the beginning of the lecture, we found that for the choice of coordinate system we made, I was diagonal. That means that the principal axes were the coordinate axes. This is what I meant by saying that we had chosen a coordinate system that wasn't stupid. In fact, we will see next time how to avoid stupid coordinate systems - and this will be the best way of actually calculating the moment of inertia tensor in all the cases we discuss in this course.

Finally, for the mathematicians in the class (and the Les Phys fans - which I hope is everybody), note that we can regard $\vec{\vec{I}}$ as a **bilinear form** - that is a machine that takes two vectors $\vec{A}$ and $\vec{B}$ into a number, via the map

$$\vec{A}, \vec{B} \rightarrow \vec{A} \cdot \vec{\vec{I}} \cdot \vec{B} \tag{38}$$

Because the matrix $\vec{\vec{I}}$ is symmetric, this is a symmetric bilinear form —

$$\vec{A} \cdot \vec{\vec{I}} \cdot \vec{B} = \vec{B} \cdot \vec{\vec{I}} \cdot \vec{A} \tag{39}$$

Finally, you can show that if the object is truly three dimensional, and not a mathematical abstraction like an infinitely thin rod or a point mass, then all the principal moments are greater than zero. This means that this is a positive definite, nondegenerate, symmetric bilinear form.

lecture 20

Topics:
  Guessing principal axes - theorems
  Another example - rectangular solids
  Rotation about a principal axis
  When principal moments are equal
  Proof of the relection theorem
  Finishing the impulse problem
  The final velocities
  After the impulse

Where are we now

Having described the moment of inertia tensor in the last lecture, we are now going to start to see what to do with it. First, let us sum up the most important properties.

1. The angular momentum $\vec{L}$ of a rigid body is related to its angular velocity vector $\vec{\omega}$ by the matrix equation

$$\vec{L} = \vec{\vec{I}} \cdot \vec{\omega} \tag{1}$$

   where $\vec{\vec{I}}$ is the moment of inertia tensor which in matrix form looks like

$$\vec{\vec{I}} = \begin{pmatrix} B_{yy} + B_{zz} & -B_{xy} & -B_{xz} \\ -B_{yx} & B_{xx} + B_{zz} & -B_{yz} \\ -B_{zx} & -B_{zy} & B_{xx} + B_{yy} \end{pmatrix} \tag{2}$$

   where the symmetric matrix $\vec{\vec{B}}$ is defined as

$$\vec{\vec{B}} = \sum_j m_j \, (\vec{r}_j - \vec{R}\,) \, (\vec{r}_j - \vec{R}\,) \tag{3}$$

   This is physics. It follows from the definition of the angular momentum and the equation for the motion of the parts of a rotating rigid body.

2. At any given time, there are three perpendicular axes

$$\hat{e}_j \quad \text{for } j = 1 \text{ to } 3 \tag{4}$$

   called the principal axes such that if $\vec{\omega}$ is in the $\hat{e}_j$ direction, then $\vec{L}$ is also —

$$\vec{\vec{I}} \cdot \omega\hat{e}_j = I_j \, \omega\hat{e}_j \tag{5}$$

   We can write the tensor $\vec{\vec{I}}$ in any coordinate system as a sum over the three principal axes,

$$\vec{\vec{I}} = I_1 \, \hat{e}_1 \, \hat{e}_1 + I_2 \, \hat{e}_2 \, \hat{e}_2 + I_3 \, \hat{e}_3 \, \hat{e}_3 \tag{6}$$

If we choose a coordinate system in which $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$ are the basis vectors, then $\vec{\vec{I}}$ is diagonal.

$$\vec{\vec{I}} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{7}$$

The moments I_j around the principal axes are called the principal moments. The directions of the three principal axes completely specify the orientation of the body. The principal axes in our example above are the x, y and z axes.

This is mathematics. The principal axes $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$ are the eigenvectors of the real symmetric matrix $\vec{\vec{I}}$, and the principal moments I_1, I_2 and I_3 are the corresponding eigenvalues.

3. The principal moment I_j is the moment of inertia about the principal axis $\hat{e}_j$.

4. The principal axes are tied to the body — they rotate as the body rotates. This is one of the things that makes rigid body rotations complicated if $\vec{\omega}$ is not along a principal axis.

Let us now return to our auto mechanic checking your tires for dynamic imbalance.

Static Imbalance:
Occurs when there is a heavy or light spot in the tire so that the tire won't roll evenly and the tire

Dynamic Imbalance:
Occurs when there is unequal weight on one or both sides of the tire/wheel assembly's

What the auto mechanics is trying to do is to balance your tire so that the axle is a principal axis of the tire. Once that is done, the angular momentum of the tire spinning around its axle is along the axle and does not change as the tire spins. Therefore no torque is required to keep the tire spinning on it axle, and there is no shimmy!

Guessing principal axes - theorems

Now that we know about principal axes, it is useful to discuss the question of finding the moment of inertia tensor again. The general formula in (2) and (3) is certainly something you should know, but with any luck you will seldom have to use it. Instead, you can remember a couple of more general things. The first thing to notice is that it is MUCH easier to find the moment of intertia tensor if you already know what the principal axes are. My recommendation is that instead of spending your time understanding the ins and outs of the general formula,

1. you should understand how to guess the principal axes, and

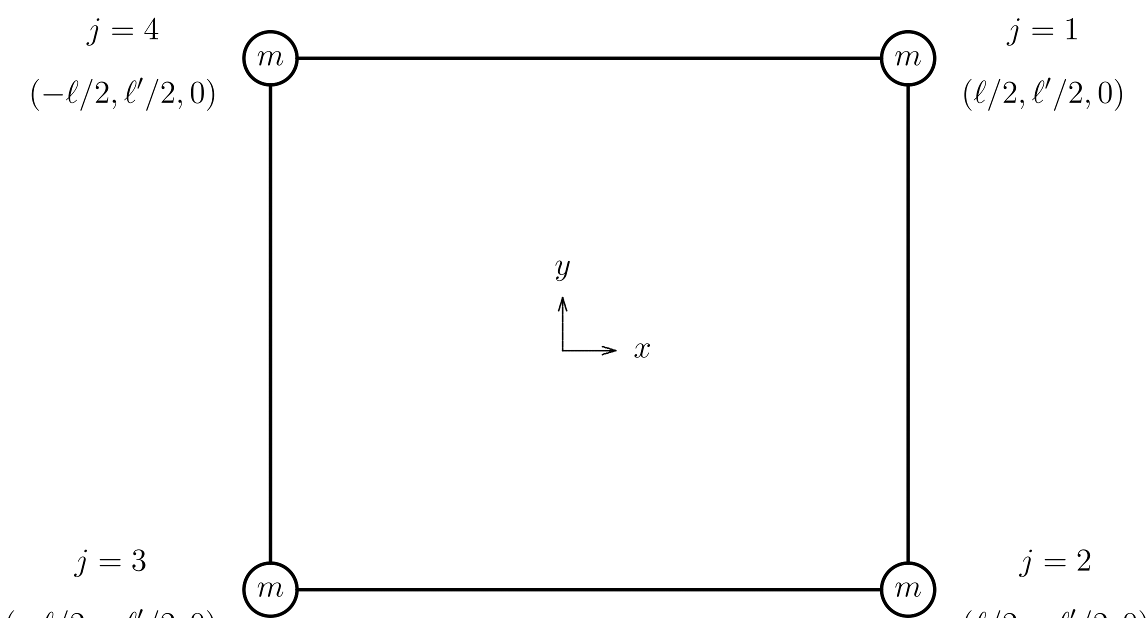
2. you should understand how to find the tensor if you already know the principal axes.

This is a far better way, in practice, to actually compute $\vec{\vec{I}}$.

Let's look at step 2 first. Suppose that you know the principal axes. Then you can rotate to a coordinate system in which the principal axes are the coordinate axes, $\hat{e}_1 = \hat{x}$, $\hat{e}_2 = \hat{y}$, $\hat{e}_3 = \hat{z}$. In this coordinate system, $\vec{\vec{I}}$ is diagonal and

$$\vec{\vec{I}}_{xx} = I_1\,, \quad \vec{\vec{I}}_{yy} = I_2\,, \quad \vec{\vec{I}}_{zz} = I_3\,, \tag{8}$$

So for example, in our rectagular example,



(9)

if we know that the x, y and z axes are principal, we can immediately conclude that $\vec{\vec{I}}_{xx} = m\,\ell'^2$ because there are four masses each of mass m and each a distance $\ell'/2$ from the x axis.

Now let's go back to step 1. The primary tool here (it will probably not surprise you) is symmetry. There are three theorems that will be useful.

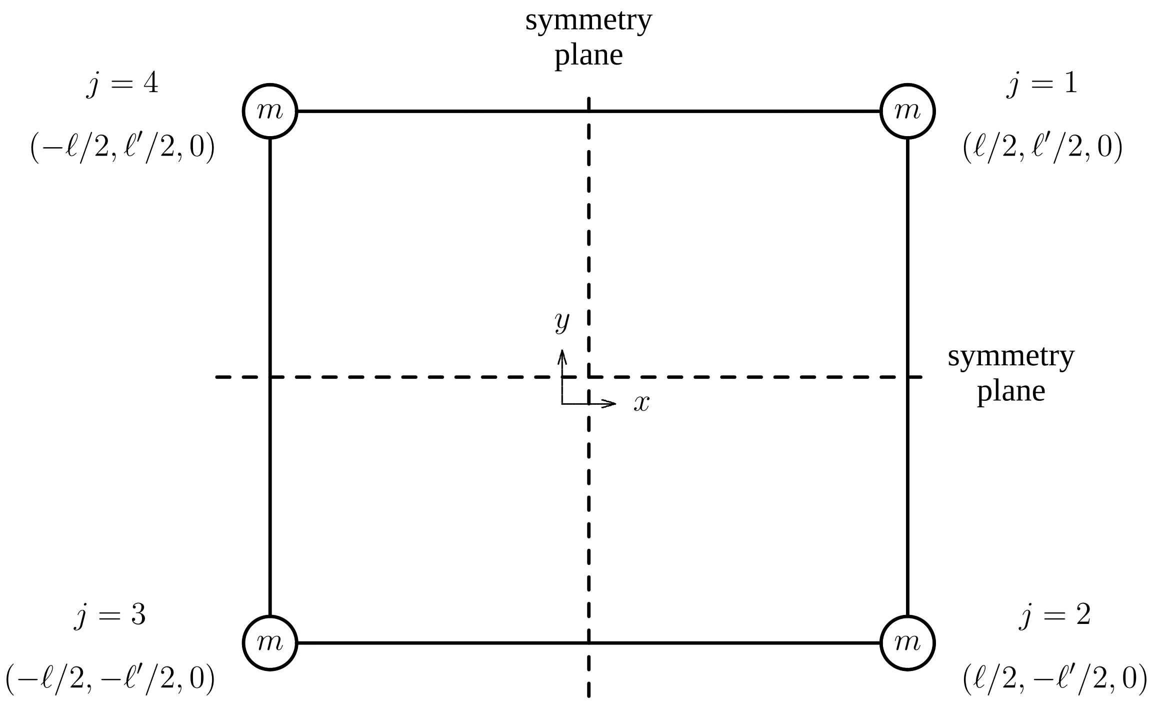
The equal-moment theorem: **Any vector in the plane formed by two different principal axes of a rigid body with equal moments is also a principal axis with the same moment.**

The reflection theorem: **If a rigid body is invariant under reflection in a plane, the vector perpendicular to this plane is a principal axis.**

The rotation theorem: **If a rigid body is invariant under a rotation of $\alpha < 2\pi$ about an axis $\hat{n}$, then (1) $\hat{n}$ is a principal axis, and (2) if $\alpha \neq \pi$, all vectors in the plane perpendicular to $\hat{n}$ are principal axes with the same principal moment.**

We will prove these later. Let's first see how to use the reflection theorem to guess the principal axes of our frame example.

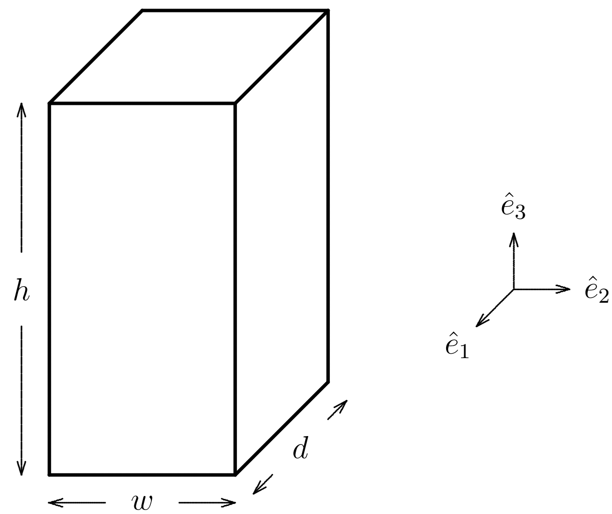
This theorem immediately implies that the principal axes in our example problem are the x, y and z axes, because the system has three planes of symmetry, the x-z and y-z planes (shown below) and the x-y plane (the plane of the paper)



(10)

Another example - rectangular solids

Consider a rectangular solid with height h, width w and depth d, with a uniform mass density,



(11)

The principal axes are the obvious ones — parallel to the edges, perpendicular to the faces. Again the reason is symmetry. The body is symmetrical under reflections through planes through the center of the body, parallel to any of the faces. So by our theorem, the vectors perpendicular to the faces are principal axes.

Now we can calculate the principal moments by calculating the moments of inertia about the principal axes . For example, moment of inertia about the axis $\hat{e}_1$ can be written as an integral

$$\int_{-d/2}^{d/2} dx \int_{-w/2}^{w/2} dy \int_{-h/2}^{h/2} dz\, \rho\,(y^2+z^2) = \frac{\rho\, dwh}{12}\,(w^2+h^2) = \frac{m}{12}\,(w^2+h^2) \tag{12}$$

Rotation about a principal axis

If a rigid body rotates freely about one of its principal axes, the angular momentum is in the same direction as $\vec{\omega}$. This makes it easy to understand the physics. Suppose that the principal axis $\hat{e}_1$ is lined up with the $\hat{z}$ direction, and $\vec{\omega}$ is also in the $\hat{z}$. Then the angular momentum is also in the $\hat{z}$ direction and is given by

$$\vec{L} = I_1\,\vec{\omega} = I_1\,|\vec{\omega}|\,\hat{z} \tag{13}$$

Conservation of angular momentum then tells you that ω remains constant as well. So the body just rotates with constant angular velocity about its principal axis, which remains in the same direction in the space frame. The other two principal axes, of course, are not fixed in the space frame. They rotate around with the body. This is illustrated in the *Mathematica* file rectangle.nb, which shows physically sensible motions of a rectangular rigid body. The blue line in the animation represents the angular velocity vector and the angular momentum vector, which are in the same direction. No problem!

But if at some time the angular velocity is not lined up with any of the principal axes, then the situation is much more complicated. Now the fact that the angular momentum is constant doesn't immediately tell you about the angular velocity, because the components of the moment of inertia tensor $\vec{\vec{I}}$ are changing with time in the space frame because of the rotation. But that in turn means that $\vec{\omega}$ is changing with time. Later, we will try to understand this much more complicated motion for a symmetric top.

Before we get started on this difficult path, I want to show you some allowed motions of rigid bodies that LOOK complicated, but actually are not. We will do that in the next section.

When principal moments are equal

Something interesting happens when two of the principal moments are equal, as in a rectangle with $w = d \neq h$. In this case,

$$I_1 = I_2 = \frac{1}{12}(h^2 + d^2) = \frac{1}{12}(h^2 + w^2) \quad \text{and} \quad I_3 = \frac{1}{12}(d^2 + w^2) = \frac{1}{6}d^2 \tag{14}$$

In this case, the principal axis corresponding to the unequal moment is unique, but the body is happy to rotate about any axis in the plane formed by the two principal axes corresponding to equal moments. The reason is the equal moment theorem: **Any vector in the plane formed by two different principal axes of a rigid body with equal moments is also a principal axis with the same moment.** Let's prove this. It follows because of the linearity of the fundamental equation that defines the principal moments. Suppose for example that $I_1 = I_2 = I$. Then

$$\vec{\vec{I}} \cdot \hat{e}_1 = I\,\hat{e}_1 \quad \text{and} \quad \vec{\vec{I}} \cdot \hat{e}_2 = I\,\hat{e}_2\,. \tag{15}$$

but that means that any linear combination of $\hat{e}_1$ and $\hat{e}_2$ satisfies the same equation:

$$\vec{\vec{I}} \cdot (a_1\hat{e}_1 + a_2\hat{e}_2) = a_1\,\vec{\vec{I}} \cdot \hat{e}_1 + a_2\,\vec{\vec{I}} \cdot \hat{e}_2 = a_1\,I\,\hat{e}_1 + a_2\,I\,\hat{e}_2 = I\,(a_1\hat{e}_1 + a_2\hat{e}_2),. \tag{16}$$

But with an arbitrary linear combination, we can get any vector in the plane spanned by $\hat{e}_1$ and $\hat{e}_2$. This completes the proof of the equal moment theorem.

Notice that the same linearity argument, (16), works even if the principal axes $\hat{e}_1$ and $\hat{e}_2$ are not orthogonal, that is $\hat{e}_1 \cdot \hat{e}_2 \neq 0$, so long as the two vectors are not equal. Indeed, the only time one can have principal axes that are not orthogonal is when they both have the same principal moment, because only then can we take linear combinations using (16) to make orthogonal axes and thus satisfy our general theorem that the momentum of inertial tensor can be made diagonal by going to the right coordinate system. Actually, (16) shows that in this case, the coordinate system is not unique. Any orthonormal pair of vectors in the plane formed by $\hat{e}_1$ and $\hat{e}_2$ is a perfectly good basis and leads to the same form for $\vec{\vec{I}}$.

A corolary of this and the general fact that every body has three perpendicular principal axes is that if you find two principal axes that are not perpendicular to one another, you know that their principal moments must be equal even before you calculate them.

The rotation theorem is closely related to this. **If a rigid body is invariant under a rotation of $\alpha < 2\pi$ about an axis $\hat{n}$, then (1) $\hat{n}$ is a principal axis, and (2) if $\alpha \neq \pi$, all vectors in the plane perpendicular to $\hat{n}$ are principal axes with the same principal moment.**

Let's prove (1) first.

First note that there must be some principal axis $\hat{e}_1$ with a nonzero component in the $\hat{n}$ directions. Then we can write

$$\hat{e}_1 = a\,\hat{n} + b\,\hat{p} \tag{17}$$

where $\hat{p}$ is some unit vector in the plane perpendicular to $\hat{n}$. Under the rotation by α about $\hat{n}$, $\hat{n}$ doesn't change, but $\hat{p}$ rotates in the plane perpendicular to $\hat{n}$. Because of the symmetry, each such rotation must give a principal axis with the same moment as $\hat{e}_1$. Then there are three possibilities:

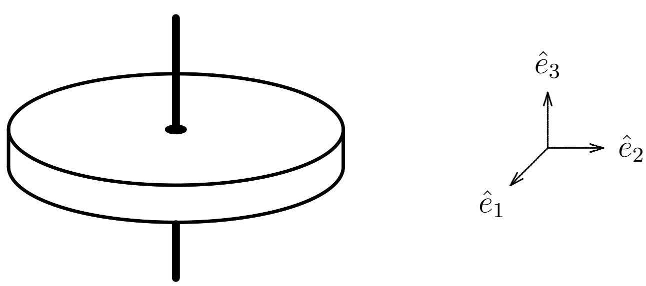
1. $b = 0$, in which case $\hat{e}_1 \propto \hat{n}$ and (1) is true;

2. $b \neq 0$ and $\alpha = \pi$, in which case $\hat{p}$ changes sign under the rotation so that $\hat{e}_1 = a\,\hat{n} + b\,\hat{p}$ and $\hat{e}_1 = a\,\hat{n} - b\,\hat{p}$ are both principal axes with the same moment - but the $\hat{e}_1$ is in the plane of two principal axes with the same moment and is a principle axis, and again (1) is true;

3. $b \neq 0$ and $\alpha = \pi$, in which case a seqence of rotations gives at least three linearly independent principal axes with the same moment, so any vector is a principal axis, and again (1) is true.

Then (2) follows because if $\alpha \neq \pi$, we can look at a principal moment perpendicular to $\hat{n}$ and it rotates into a linearly independent vector perpendicular to $\hat{n}$, and then the equal moment theorem gives the result.

A body with two equal principal moments is called a symmetric top because one simple way to get two equal moments is to have a **continuous** symmetry that rotates $\hat{e}_1$ into $\hat{e}_2$, as in a top like that shown below:



(18)

In the case of a symmetric top, it is obvious that any vector in the plane of the disk is a principal axis, because all these directions are physically equivalent. We can rotate one into another. Note, however, that as we have seen above such a symmetry is not necessary. As long as $I_1 = I_2$, any vector in the plane of $\hat{e}_1$ and $\hat{e}_2$ is a principal axis. In the rectangular solid with $d = w$, these directions don't all look equivalent, but they are equivalent so far as rotations are concerned. Thus for example a rectangular solid with $d = w$ can rotate happily about any axis perpendicular to its length. This rotation looks more complicated than the rotation about an axis perpendicular to a face, but really it isn't. An example is shown in the *Mathematica* file cylinder.nb, which illustrates that the behavior of such a rectagular solid with respect to rotations is the same as that of a cylinder.

A nice symmetrical body like a sphere has all three principal moments equal,

$$I_1 = I_2 = I_3 = I = \frac{2}{5}\, m\, r^2 \tag{19}$$

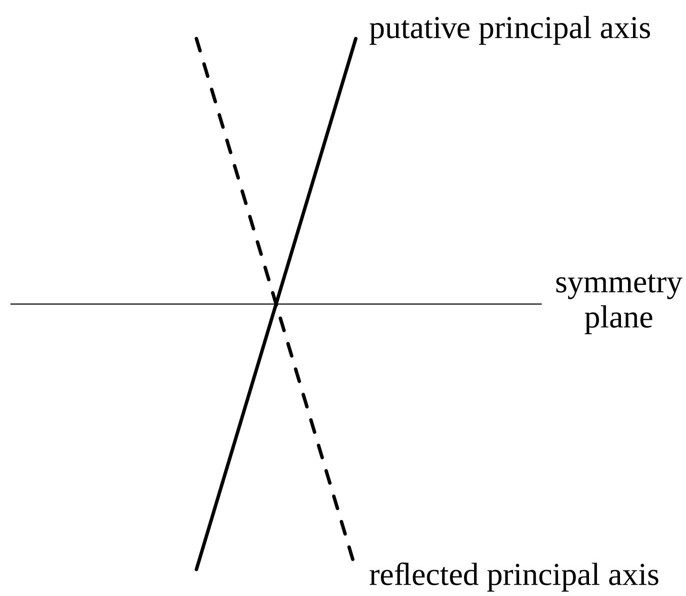
where m is the mass and r is the radius. For such a symmetrical object, the angular momentum is always in the direction of the angular velocity

$$\vec{L} = I\,\vec{\omega} \tag{20}$$

As with the symmetrical top, (20) does not depend on the body actually being rotationally symmetric. It depends only on the fact that $I_1 = I_2 = I_3$. For example, a solid cube is clearly not invariant under arbitrary rotations. However, it does have $I_1 = I_2 = I_3$ because it is symmetric under **discrete** rotations by $90°$ about lines perpendicular to the faces. This is enough to guarantee that the angular momentum is parallel to the angular velocity for a cube. So a cube is happy to rotate about any axis at constant angular velocity, whether or not the axis is lined up with one of the special axes in the cube. The same is true of any other regular solid, as in the *Mathematica* file regular-solid.nb. The three moments of inertia are equal.

Proof of the reflection theorem

Here is the theorem again. **If a rigid body is invariant under reflection in a plane, the vector perpendicular to this plane is a principal axis.** Here is the proof. Suppose we have a plane of reflection and we construct the vector perpendicular to the plane. Some principal axis must have at least some component along this vector. But then because of the symmetry, we know that if we reflect this axis in the symmetry plane, we must get another principal axis with the same principal moment, as illustrated below:



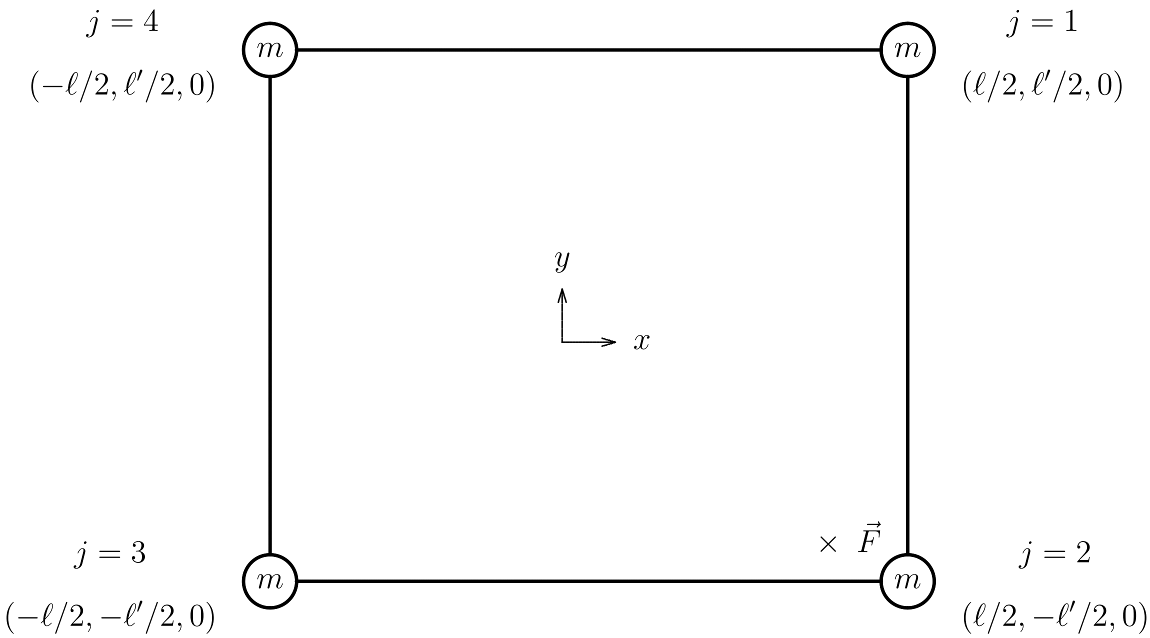
(21)

Now there are two possibilities. One possibility is that the reflected axis coincides with the original. The only way this can happen is for the original axis to be perpendicular to the symmetry plane, in which case the theorem is true. The other possibility is that the reflected axis is different. But then there are two distinct axes with the same principal moment (because they are physically equivalent due to the symmetry). But then any axis in the plane formed by these two is a perfectly good principal axis — in particular the one perpendicular to the symmetry plane — so again the theorem is true.

This is illustrated in *Mathematica* file axes2.nb. There is a slight subtlety that is also illustrated in this file. A reflection is not equivalent to a rotation by $180°$. And in fact, symmetry under a rotation by $180°$ doesn't tell you anything about principal moments in the plane of the rotation. Symmetry under a rotation by $2\pi/n$ for $n \geq 3$ is enough to guarantee that any vector in the plane perpendicular to the axis of rotation is a principal axis with a principal moment independent of the angle in the plane. But this is not true for $n = 2$ — symmetry under a rotation by π or $180°$. This only tells you that the axis is a principal moment.

Finishing the impulse problem

Let's review where we are with the light rigid rectangular frame with masses at the corners, as shown:



(22)

where the force acts for a very short time, and the impulse is into the paper, with the form

$$\int dt\, \vec{F} = -\hat{z}\, P \tag{23}$$

for positive P. Earlier, we found the center of mass to be at the origin and we found the moment of inertia tensor for this object about its center of mass to be

$$\overleftrightarrow{I}_{xx} = m\,\ell'^2 \qquad \overleftrightarrow{I}_{yy} = m\,\ell^2 \qquad \overleftrightarrow{I}_{zz} = \overleftrightarrow{I}_{xx} + \overleftrightarrow{I}_{yy} \qquad \overleftrightarrow{I}_{xy} = \overleftrightarrow{I}_{xz} = \overleftrightarrow{I}_{yz} = 0 \tag{24}$$

To complete the story, we need to compute the velocity of the center of mass, and the angular velocity about the center of mass. The velocity is easy. We know the impulse, and we know the total mass, so we can compute the change in momentum and thus the final velocity. We'll do this in a minute. First we will do the interesting part and compute the angular velocity.

To do this, we first compute the impulsive torque about the center of mass — that is

$$\int dt\, \vec{\tau} \tag{25}$$

over the short period during which the force acts. This is

$$\int dt\, (\vec{r}_2 - \vec{R}) \times \vec{F} = (\ell/2, -\ell'/2, 0) \times (0, 0, -P) = (P\ell'/2, P\ell/2, 0) \tag{26}$$

where we have used the fact that the force acts for a very short time, so that $\vec{r}_2$ and $\vec{R}$ do not change significantly while the force acts, and they are therefore effectively constants and can be taken outside the integral in (26). This is equal to the change in angular momentum about the center of mass, which is the final angular momentum about the center of mass.

$$\vec{L} = (P\ell'/2, P\ell/2, 0) \tag{27}$$

But the angular momentum is related to the angular velocity by

$$\vec{L} = \vec{\vec{I}} \cdot \vec{\omega} \tag{28}$$

Because $\vec{L}$ is conserved after the force acts, (28) remains true for any time afterwards. But here is the point. Immediately after the force acts, the body has still not had a chance to move from its position in (22). Therefore, the moment of inertial tensor is given by (24), which implies (using (24) and (27))

$$(P\ell'/2, P\ell/2, 0) = (m\ell'^2\, \omega_x, m\ell^2\, \omega_y, \vec{\vec{I}}_{zz}\omega_z) \tag{29}$$

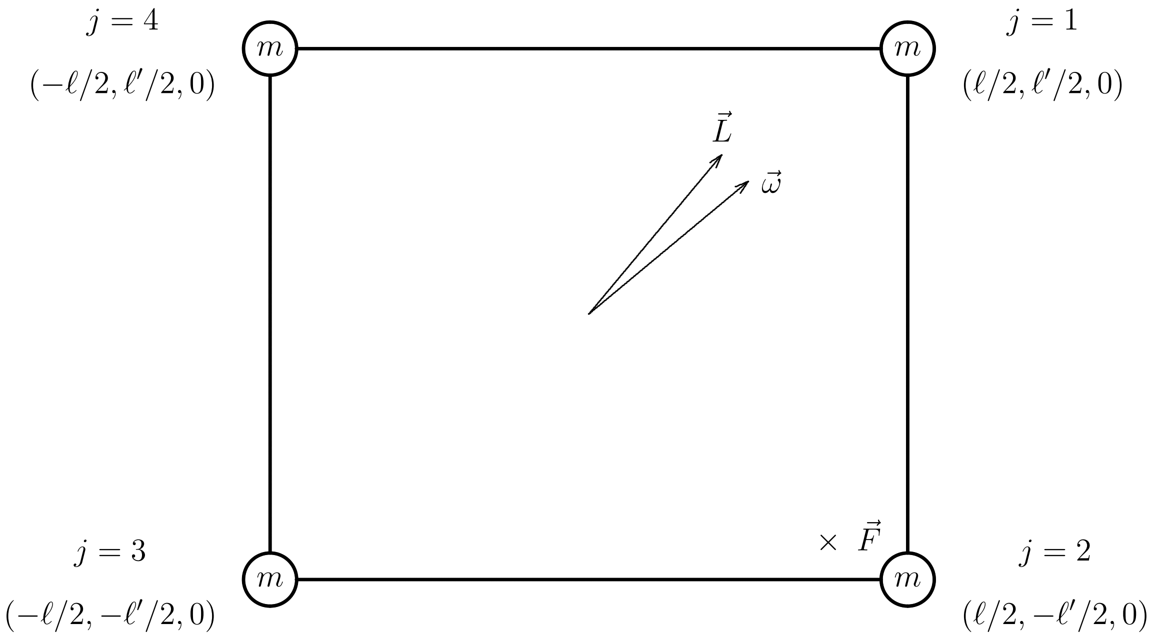
which in turn implies

$$\omega_x = \frac{P}{2m\ell'} \qquad \omega_y = \frac{P}{2m\ell} \qquad \omega_z = 0 \tag{30}$$

This is the desired result for the angular velocity about the center of mass immediately after the hammer blow:

$$\vec{\omega} = \left(\frac{P}{2m\,\ell'}, \frac{P}{2m\,\ell}, 0 \right) \tag{31}$$

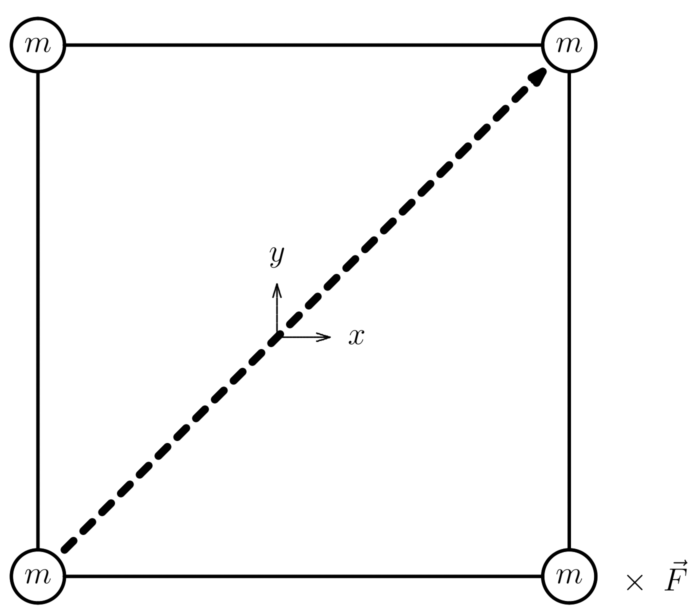
The directions of the vectors $\vec{L}$ and $\vec{\omega}$ are shown below (in an arbitrary relative normalization, because they have different units).



(32)

It is amusing (though not particularly signficant) that $\vec{\omega}$ points at mass 1.

(31) is very easy to understand in the limit $\ell = \ell'$. In this case, the frame is a square, the x and y principal moments are equal, so we can take the principal axes to be the diagonals, and (31) simply describes rotation about the obvious diagonal, shown as the dashed line below:



(33)

The final velocities

Now to get the velocities of the individual masses, we use the fact that the motion of the body is a combination of the translational motion of the center of mass plus the rotational motion about the center of mass. The final velocity of the center of mass is the impulse divided by the total mass:

$$\vec{V} = \left(0, 0, -\frac{P}{4m}\right) \tag{34}$$

The velocity of mass j is then given by

$$\vec{V} + \vec{\omega} \times (\vec{r}_j - \vec{R}) \tag{35}$$

where $\vec{\omega}$ is given by (31). Here they are:

$$\begin{aligned}\vec{v}_1 &= \left(0, 0, -\frac{P}{4m}\right) + \left(\frac{P}{2m\,\ell'}, \frac{P}{2m\,\ell}, 0\right) \times \left(\frac{\ell}{2}, \frac{\ell'}{2}, 0\right) \\ &= \left(0, 0, -\frac{P}{4m}\right)\end{aligned} \tag{36}$$

$$\begin{aligned}\vec{v}_2 &= \left(0, 0, -\frac{P}{4m}\right) + \left(\frac{P}{2m\,\ell'}, \frac{P}{2m\,\ell}, 0\right) \times \left(\frac{\ell}{2}, -\frac{\ell'}{2}, 0\right) \\ &= \left(0, 0, -\frac{3P}{4m}\right)\end{aligned} \tag{37}$$

$$\begin{aligned}\vec{v}_3 &= \left(0, 0, -\frac{P}{4m}\right) + \left(\frac{P}{2m\,\ell'}, \frac{P}{2m\,\ell}, 0\right) \times \left(-\frac{\ell}{2}, -\frac{\ell'}{2}, 0\right) \\ &= \left(0, 0, -\frac{P}{4m}\right)\end{aligned} \tag{38}$$

$$\begin{aligned}\vec{v}_4 &= \left(0, 0, -\frac{P}{4m}\right) + \left(\frac{P}{2m\,\ell'}, \frac{P}{2m\,\ell}, 0\right) \times \left(-\frac{\ell}{2}, \frac{\ell'}{2}, 0\right) \\ &= \left(0, 0, \frac{P}{4m}\right)\end{aligned} \tag{39}$$

After the impulse

I keep saying that things get so complicated after the impulse that I don't want to require you to understand the general case. But I think it may be useful to see how we would use *Mathematica* to study it. *Mathematica* allows us to follow the motion numerically in a pretty simple way, because it gives us simple tools for dealing with rotations. Let's look at the notebook after-the-impulse.nb.

The key is to keep track of the rotation matrix $\vec{\vec{R}}(t)$ (called r in the notebook) that rotates the body from its initial orientation to its orientation at time t. If we know this, and we know $\vec{\omega}(t)$ (called ω in the notebook), then we can calculate $\vec{\vec{R}}(t+dt)$ —

$$\vec{\vec{R}}(t+dt) = \text{RotationMatrix}[|\vec{\omega}|\,dt, \vec{\omega}] \cdot \vec{\vec{R}}(t) \tag{40}$$

where RotationMatrix[θ,ω] (abbreviated as rm in the notebook) rotates by θ about the vector $\vec{\omega}$. This is just translation into matrix language of our cross product formula for the motion of a vector rotating with angular velocity $\vec{\omega}$. But we can also find $\vec{\omega}$ from the fixed angular momentum $\vec{L}$ because we can find the moment of inertia tensor in terms of $\vec{\vec{R}}(t)$,

$$\vec{\vec{I}}(t) = \vec{\vec{R}}(t) \cdot \vec{\vec{I}}(0) \cdot \vec{\vec{R}}(t)^T \tag{41}$$

Then

$$\vec{\omega}(t) = \vec{\vec{I}}(t)^{-1} \cdot \vec{L} \tag{42}$$

That is how the animation works. We look at the motion after the impulse in a frame moving along with the center of mass. The green line shows $\vec{\omega}$ and the red line shows $\vec{L}$. AfterIt wanders all over the place as the body tumbles out of control.

lecture 20

Topics:

The free symmetric top in the space frame
Vectors in the body frame
The free symmetric top in the body frame
Euler's equations
The free symmetric top ala Euler's
The tennis racket theorem

As you know, a spinning top subject to a gravitational torque will precess. In fact, the situation is a bit more complicated. If the top doesn't get started quite right, the motion is more complicated. The precession of the symmetry axis about the direction of the angular momentum that we saw in our discussion of the free symmetric top gives wise to the phenomenon of nutation in precessing symmetric tops. We will see how this works, and then go on to see what we can say about the general case by looking at it in the body frame.

The free symmetric top in the space frame

Now we are finally going to face the music and try to figure out analytically what happens AFTER the initial impulse when the angular momentum and the angular velocity are not in the same direction. However, I am not going to be so cruel as to try to completely solve the problem in the case of a general rigid body. As I said, that is too hard. But we will analyze completely only the case of the free precession of a symmetric top, with two equal moments of inertia. It not just that I am taking pity on you. I think that it is possible to understand this in a way that is probably impossible for the general case. Since I am interested in getting you to really understand things, rather than just cramming facts into your head, I will focus on the simpler problem. I am going to do this in a very slick way. We won't need any complicated differential equations — just the relation $\vec{L} = \vec{\vec{I}} \cdot \vec{\omega}$ and the connection of the cross product to uniform circular motion. I hope that you will find it so simple that you will be able easily to reproduce it yourself, though you may find yourself wondering what happened.

Later, we will go back and do it the hard way.

Suppose that a rigid body with two equal principal moments (this makes it a symmetric top) is floating out in space, and is rotating freely about its center of mass. Because there are no external torques on the system, the angular momentum $\vec{L}$ is conserved. If $\vec{L}$ points along a principal axis of the body, then the angular velocity vector $\vec{\omega}$ is parallel to $\vec{L}$ and it also remains constant. The interesting mathematical question is what happens to $\vec{\omega}$ when it is not parallel to $\vec{L}$. The physical question is then what the resulting motion looks like!

You should not be surprised that the key is the connection between the angular momentum and the angular velocity through the moment of inertia tensor —

$$\vec{L} = \vec{\vec{I}} \cdot \vec{\omega} \tag{1}$$

The important point is that in the space frame, the principal axes of the top, $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$, are all moving around as the body rotates, but at any given time, they form a complete orthonormal set of vectors in terms of which we can expand any vector. For example

$$\vec{\omega} = \hat{e}_1 \; (\hat{e}_1 \cdot \vec{\omega}) + \hat{e}_2 \; (\hat{e}_2 \cdot \vec{\omega}) + \hat{e}_3 \; (\hat{e}_3 \cdot \vec{\omega}) = \omega_1 \, \hat{e}_1 + \omega_2 \, \hat{e}_2 + \omega_3 \, \hat{e}_3 \tag{2}$$

which just says that $\vec{\omega}$ can be expanded in the basis vectors, $\hat{e}_j$. Using (2) in this basis, (1) becomes

$$\vec{L} = I_1 \, \omega_1 \, \hat{e}_1 + I_2 \, \omega_2 \, \hat{e}_2 + I_3 \, \omega_3 \, \hat{e}_3 \tag{3}$$

The relation (3) is entirely general, valid for an arbitrary moment of inertia, but here, we are dealing with a symmetric top, for which two of the principal moments are equal. Let's assume that the two equal moments are I_1 and I_2, with corresponding principal axes $\hat{e}_1$ and $\hat{e}_2$. As you know from the equal-moment theorem, this means that any axis in the plane of $\hat{e}_1$ and $\hat{e}_2$ (which is the plane perpendicular to $\hat{e}_3$) is a principal axis with the same principal moment. To remind you of this, I will define $I_\perp$ as

$$I_\perp = I_1 = I_2 \tag{4}$$

Then we can write (3) as

$$\vec{L} = I_\perp \, \omega_1 \, \hat{e}_1 + I_\perp \, \omega_2 \, \hat{e}_2 + I_3 \, \omega_3 \, \hat{e}_3 \tag{5}$$

Another way of deriving (5) is to use our dyadic expression for the moment of inertia tensor,

$$\vec{\vec{I}} = I_1 \, \hat{e}_1 \hat{e}_1 + I_2 \, \hat{e}_2 \hat{e}_2 + I_3 \, \hat{e}_3 \hat{e}_3 \tag{6}$$

which for a free symmetric top looks like

$$\vec{\vec{I}} = I_1 \, \hat{e}_1 \hat{e}_1 + I_\perp \, \hat{e}_2 \hat{e}_2 + I_\perp \, \hat{e}_3 \hat{e}_3 \tag{7}$$

Now because of (4), we can eliminate all reference to $\hat{e}_1$ and $\hat{e}_2$ which enormously simplifies the problem. If we now multiply both sides of (2) by $I_\perp$ and subtract both sides of the result from (5), we get

$$\begin{array}{rc} & \vec{L} = I_\perp \, \omega_1 \, \hat{e}_1 + I_\perp \, \omega_2 \, \hat{e}_2 + I_3 \, \omega_3 \, \hat{e}_3 \\ - & \left[\quad I_\perp \, \vec{\omega} = I_\perp \, \omega_1 \, \hat{e}_1 + I_\perp \, \omega_2 \, \hat{e}_2 + I_\perp \, \omega_3 \, \hat{e}_3 \quad \right] \\ \hline & \vec{L} - I_\perp \, \vec{\omega} = (I_3 - I_\perp) \, \omega_3 \, \hat{e}_3 \end{array} \tag{8}$$

or

$$\vec{L} = I_\perp \, \vec{\omega} + (I_3 - I_\perp) \, \omega_3 \, \hat{e}_3 = I_\perp \, \vec{\omega} + I_\perp \, \Omega \, \hat{e}_3 \tag{9}$$

where

$$\Omega = \frac{(I_3 - I_\perp) \, \omega_3}{I_\perp} \tag{10}$$

There is an ambiguity in the sign of $\hat{e}_3$ because the symmetry axis is not directional, and we resolve this by always taking $\omega_3 > 0$. The constant Ω has units of angular velocity, and it always satisfies the inequality

$$-\omega_3 \leq \Omega \leq \omega_3 \tag{11}$$

because

$$-I_\perp \leq I_3 - I_\perp \leq I_\perp \tag{12}$$

We will talk more about this in a general way soon.

It will sometimes be convenient to rewrite (9) as an equation for ω,

$$\vec{\omega} = \vec{L}/I_\perp - \Omega\hat{e}_3 \tag{13}$$

The relations (9) and (13) have nothing in them beyond the connection between angular momentum and angular velocity and the fact that $I_1 = I_2 = I_\perp$. They are true at any time, but do not tell us directly about time dependence. However, it is the simplicity of these relations that allows us to solve this problem much more easily than we could for a general rigid body. Mathematically, it is the equality of I_1 and I_2 that allows us to eliminate both $\hat{e}_1$ and $\hat{e}_2$ from (9) and (13). Physically, this is particularly important because the one remaining principal axis in (9) and (13) is the one we see in the typical symmetric top. The symmetry axis is special, and that makes (9) and (13) particularly useful.

We can now get the time dependence by judiciously using the connection between rotation and the cross product.

The first step is to notice that $\hat{e}_3$ is a principal axis fixed in the body, and it is therefore rotating with the body. Like any other vector fixed in the body, $\hat{e}_3$ has a time dependence given the the cross product with the angular velocity vector $\vec{\omega}$.

$$\frac{d}{dt}\hat{e}_3 = \vec{\omega} \times \hat{e}_3 \tag{14}$$

The relation (14) is a differential equation for the time dependence of $\hat{e}_3$, but it is not particularly useful, because $\vec{\omega}$ is also changing with time in a way that we do not yet know. Instead of trying to bull our way through this, we can use (13) to write

$$\frac{d}{dt}\hat{e}_3 = \left(\vec{L}/I_\perp - \Omega\hat{e}_3\right) \times \hat{e}_3 = \frac{\vec{L}}{I_\perp} \times \hat{e}_3 \tag{15}$$

This is a much more useful equation, because $\vec{L}$ is constant. In fact, we don't even have to solve it, because we already know exactly what it means geometrically. The time dependence of $\hat{e}_3$ given in (15) is what we expect if $\hat{e}_3$ is rotating about an axis in the direction of $\vec{L}$ with angular velocity $\vec{L}/I_\perp$. Since $\vec{L}$ is constant in the space frame, this must be what is actually going on for all times — $\hat{e}_3$ is undergoing uniform circular motion with angular velocity $\vec{L}/I_\perp$. Now we are essentially done. From this, we can calculate everything.

First notice that the component of the angular momentum in the $\hat{e}_3$ direction (along the symmetry axis of the top), $\vec{L} \cdot \hat{e}_3$ is constant. Geometrically, this is because $\hat{e}_3$ is rotating around the fixed vector $\vec{L}$, so while its component perpendicular to $\vec{L}$ is constantly changing, its parallel component is not. Analytically it follows from (15) as follows:

$$\frac{d}{dt}\left(\vec{L}\cdot\hat{e}_3\right) = \vec{L}\cdot\frac{d}{dt}\hat{e}_3 = \vec{L}\cdot\left(\frac{\vec{L}}{I_\perp}\times\hat{e}_3\right) = 0 \tag{16}$$

But $\vec{L}\cdot\hat{e}_3$ is related to ω_3 by (5),

$$\vec{L}\cdot\hat{e}_3 = I_3\omega_3 \tag{17}$$

Thus (16) means that ω_3 is constant,

$$\frac{d}{dt}\omega_3 = 0 \tag{18}$$

This in turn means that Ω is constant also, so the coefficients in (9) and (13) are constant in time. Note also that using (17), Ω can be written as

$$\Omega = \frac{|\vec{L}\,|}{I_\perp}\frac{I_3 - I_\perp}{I_3}\cos\theta \tag{19}$$

where θ is the angle between $\vec{L}$ and the symmetry axis (between 0 and $\pi/2$).

Because of (13) and (18), we know that $\vec{\omega}$ is a fixed linear combination of $\vec{L}$ and $\hat{e}_3$, and therefore since $\hat{e}_3$ is undergoing uniform circular motion with angular velocity $\vec{L}/I_\perp$, then $\vec{\omega}$ is also. We can see this geometrically because (9) or (13) and (16) imply that $\vec{L}$, $\hat{e}_3$ and $\vec{\omega}$ are fixed in a plane and then from (15) we can conclude that the plane formed by $\vec{L}$, $\hat{e}_3$ and $\vec{\omega}$ rotates about the fixed angular momentum with angular velocity $\vec{L}/I_\perp$. We can also see this analytically

$$\frac{d}{dt}\vec{\omega} \tag{20}$$

$$= \frac{d}{dt}\left(\vec{L}/I_\perp - \Omega\hat{e}_3\right) \tag{21}$$

$$= \frac{d}{dt}\left(-\Omega\hat{e}_3\right) \tag{22}$$

$$= \frac{\vec{L}}{I_\perp}\times\left(-\Omega\hat{e}_3\right) \tag{23}$$

$$= \frac{\vec{L}}{I_\perp}\times\left(\vec{L}/I_\perp - \Omega\hat{e}_3\right) \tag{24}$$

$$= \frac{\vec{L}}{I_\perp}\times\vec{\omega} \tag{25}$$

Notice that we have still not figured out explicitly what the perpendicular axes $\hat{e}_1$ and $\hat{e}_2$ are doing. But this information is implicit in the relation (13). In words, this relation together with (15) says that the angular velocity of the top can be taken apart into two components:

1. the rotation of the top around its symmetry axis $\hat{e}_3$ with angular velocity $-\Omega\,\hat{e}_3$; and

2. the rotation of $\hat{e}_3$ around the fixed angular momentum with angular velocity $\vec{L}/I_\perp$.

Let's review the few key steps in this derivation. First we used the form of the moment of inertia tensor to eliminate

Let us now look at the result in animated form in the *Mathematica* worksheet freetop.nb This is an animation of a rotating rectangular solid, with two opposite sides square, so that it has two equal principal moments. This is an allowed free rotation of the body in space. The angular momentum is $\vec{L} = L\,\hat{z}$ and is constant. The ratio of the length in the $\hat{e}^3$ direction to that in the $\hat{e}_{1,2}$ directions starts at 2 with $I_3/I_\perp = 0.4$ less than one, corresponding to a prolate object — long and thin along the symmetry axis. The angle θ is the angle between the angular momentum (the $\hat{z}$ axis) and $\hat{e}_3$, the symmetry axis of the body. The initial θ is zero. Then $\vec{\omega} = \hat{z}\,L/I_3$.

Now work your way towards larger initial angles. If you switch from "solid" to "frame" you see three lines representing vectors in addition to to rotating solid. The green line is the angular velocity vector, $\vec{\omega}$. The vertical purple line and the red line are the components of the $\vec{\omega}$ vector along the directions of the angular momentum $\vec{L}$ and the symmetry axis $\hat{e}_3$. Thus these vectors are a visual representation of the fundamental relation between $\vec{L}$, $\vec{\omega}$ and $\hat{e}_3$, expressed in (9) and (13).

I hope that these animations will help you to understand more deeply the connection between (13) and (15). The rate of rotation of the plane is always $|\vec{L}|/I_\perp$, which is fixed in the animations. One way of describing the crucial relation (13) is that we have taken the angular velocity vector apart into a component along $\hat{e}_3$ that describes the rotation of the object around its symmetry axis and the vertical component $\vec{L}/I_\perp$ that describes the rotation of the plane of the symmetry axis. Writing $\vec{\omega}$ this way is great, because you should be able to see both components physically. Notice that for the prolate object, for an initial angle less than $\pi/2$, $\vec{\omega}$ is longer than $\vec{L}/I$. This is prolateness in action - the moment of inertial about the long axis is smaller than $I_\perp$, so (15) implies that the magnitude of the angular momentum must be larger than $|\vec{L}|/I_\perp$. The spin around the axis and the spin around the axis of the angular momentum add positively to the total angular velocity. For an oblate object, the opposite is true. The angular velocity vector is shorter than $|\vec{L}|/I_\perp$, which means that the spin around the axis is in the opposite direction. But in this case as well, the plane of $\hat{e}_3$ and ω rotates about the fixed angular momentum, which has the physical effect that the symmetry axis precesses. It is worth staring at this animation for a long while.

Vectors in the body frame

The derivation we gave at the beginning of the lecture of the motion of the free symmetric top in the space frame is slick and (I hope) easy to understand, but limited to the case $I_1 = I_2$. While we will not solve the more general problem completely, it is interesting to set it up mathematically in the body frame, rotating along with this rigid body. At least, in this frame, the moments of inertia do not change. We will take the origin to be the center of mass in both the space frame and the body frame. We will also take our coordinate axes in the body frame to correspond to the three principal axes of the rigid body. Then the equations that govern the motion of the body take a fairly simple form. Of course, even if we can solve these equations, it doesn't tell us everything, because it will still take some more work to understand how the body is moving in the space frame. But at least this is a start.

Our starting point will be the relation between the change of a vector with respect to the space frame and the change with respect to the rotating body frame. This is going to be important for other things as well, so we will pause to go over it slowly.

Consider a vector $\vec{A}$. We will call the rate of change of the vector with respect to the body frame

$$\frac{\delta \vec{A}}{\delta t} \,. \tag{26}$$

We have invented this funny symbolism to avoid getting this confused with the rate of change of the vector with respect to the space frame, for which we continue to use the conventional symbols,

$$\frac{d \vec{A}}{dt} \,. \tag{27}$$

By definition, if $\frac{\delta \vec{A}}{\delta t} = 0$, the vector is fixed with respect to the body frame, and then we know that its change with respect to the space frame is simply $\vec{\omega} \times \vec{A}$ where $\vec{\omega}$ is the angular velocity of the body. If $\frac{\delta \vec{A}}{\delta t} \neq 0$, then the change with respect to the space frame is the sum of this and the contribution from the body's rotation:

$$\frac{d \vec{A}}{dt} = \frac{\delta \vec{A}}{\delta t} + \vec{\omega} \times \vec{A} \,. \tag{28}$$

What I find confusing about (28) is that it seems to refer to two different frames — the space frame and the body frame. In which frame is it true? The answer is in either one! The quantities on both sides of the equation are vectors that can be expressed either in the space frame or in the body frame. The explicit coordinates of $\vec{A}$ and its derivatives and $\vec{\omega}$ will change when we go from one frame to another, but (28) is true in either. This is why I was careful to use the words "with respect to" rather than the word "in" in defining $\frac{d\vec{A}}{dt}$ and $\frac{\delta \vec{A}}{\delta t}$. These objects are the change of the vector $\vec{A}$ **with respect to** the space frame and the body frame respectively, but they can be described **in** either frame (or for that matter, any other coordinate system).

There is something very subtle and interesting going on here which I often get confused about myself. It is worth trying to say this in a few different ways because it gets into some issues that are usually not discussed in enough detail. Bear with me, because I am still trying to figure out the best way to say this. In Newtonian mechanics, we tend to use the terms "coordinate system" and "frame" interchangeably, and for example talk about a "moving coordinate system." This is potentially confusing because we get into the problems we have just been discussing. I think that a better way of saying things in nonrelativistic Newtonian mechanics is to talk about a coordinate system only as the system we use to describe the components of vectors at on particular time. A "frame" is then something that specifies what coordinate system to use at each time. If we try to do this consistently, we see that it makes no sense to talk about the time derivative of a vector without specifying the frame, because the components of the time derivative involve the difference between to components of the vector at two neighboring times which must be specified in the coordinate system that the frame tells you to use at those two different times. **Any time we write a time derivative of a vector (or tensor), we must specify a frame. However, the result of**

the differentiation in a particular frame at a particular time is a vector that has an invariant meaning, and can be described in any coordinate system at that time.

Curiously, this almost comes more naturally in relativity. If we think of the notion of "frame" more physically, as we are forced to do once velocities get close to c, we are naturally constrained to do the right things, because we have to specify **HOW** we actually measure the coordinates of our vectors. Of course this means that we need to have a coordinate system that everyone in the frame agrees on at each different time. The fact that time gets mixed up with space in relativity forced us to this more physical picture of a frame, but in fact we really need it just to keep from getting confused even at low speeds, at least if we are thinking about rotating or otherwise accelerating frames.

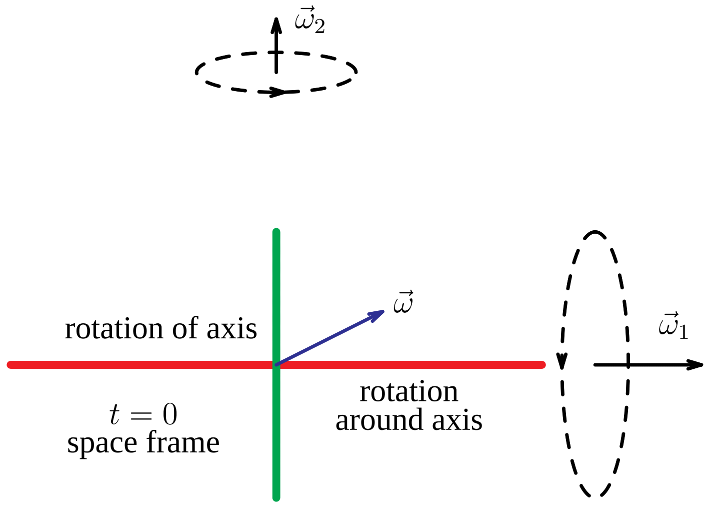
Let's discuss an example. Consider the vector $\vec{\omega}$ itself. Here (28) implies

$$\frac{d\vec{\omega}}{dt} = \frac{\delta\vec{\omega}}{\delta t} + \vec{\omega} \times \vec{\omega} = \frac{\delta\vec{\omega}}{\delta t} \,. \tag{29}$$

The rate of change of ω with respect to the body frame is the same vector as the rate of change of omega with respect to the space frame. First consider the situation in which $\vec{\omega} = \omega\hat{n}$ is constant in space. This means that the body is rotating about the fixed axis $\hat{n}$ with constant angular velocity ω. In the body frame, this axis is fixed in the body, and again $\vec{\omega}$ is constant. Thus

$$\frac{d\vec{\omega}}{dt} = 0 \Rightarrow \frac{\delta\vec{\omega}}{\delta t} = 0 \tag{30}$$

But the relation (29) may seem a bit odd in general, so let's see how it works in a nontrivial example. Considering the object you studied in problem 9.1 will give us something specific to think about and will also illustrate some of the curious features of angular velocity.



(31)

This is intended to show a rigid body made of two perpendicular rods rotating in the space frame about one of its axes $\hat{e}_1$ (the long one — red if you can see the figure in color) with angular velocity

$$\vec{\omega}_1 = \omega_1 \hat{e}_1 \tag{32}$$

which is constant in the body frame and with the axis $\hat{e}_1$ rotating about the vertical with angular velocity

$$\vec{\omega}_2 = \omega_2 \hat{z} \tag{33}$$

constant in the space frame. The total angular velocity is the sum $\vec{\omega} = \vec{\omega}_1 + \vec{\omega}_2$ and the component of the angular velocity in the direction of the axis is executing uniform circular motion and the rate of change of $\vec{\omega}$ is

$$\frac{d}{dt}\vec{\omega} = \vec{\omega}_2 \times \vec{\omega} = \vec{\omega}_2 \times (\vec{\omega}_1 + \vec{\omega}_2) = \vec{\omega}_2 \times \vec{\omega}_1 \tag{34}$$

In the space frame, as the red $\hat{e}_1$ axis rotates around with angular velocity $\vec{\omega}_2$, you see the shorter crosspiece rotating around the red $\hat{e}_1$ axis with angular velocity $\vec{\omega}_1$. For $\omega_1 = 3\Omega$ and $\omega_2 = \Omega$, this is what was shown in the *Mathematica* notebook as9-rotations.nb that you studied on the problem set.

The animation is based on the fact that our description in words of the angular velocity describes not just what ω looks like at a particular moment, but also how it evolves with time (at least if we impose an initial position). For example, if the axis $\hat{e}_1$ is in the $\hat{x}$ direction at time $t = 0$, then at an arbitrary time

$$\hat{e}_1 = \hat{x}\cos\omega_2 t + \hat{y}\sin\omega_2 t \tag{35}$$

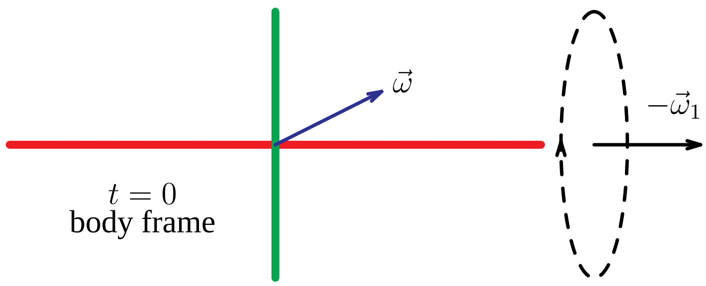
$$\vec{\omega}_1 = \omega_1\hat{x}\cos\omega_2 t + \omega_1\hat{y}\sin\omega_2 t \tag{36}$$

$$\vec{\omega} = \omega_1\hat{x}\cos\omega_2 t + \omega_1\hat{y}\sin\omega_2 t + \omega_2\hat{z} \tag{37}$$

From this and (33), you can explicitly verify (34):

$$\frac{d}{dt}\vec{\omega} = -\omega_1\omega_2\hat{x}\sin\omega_2 t + \omega_1\omega_2\hat{y}\cos\omega_2 t = \omega_2\hat{z} \times (\omega_1\hat{x}\cos\omega_2 t + \omega_1\hat{y}\sin\omega_2 t) \tag{38}$$

Now what does this look like in the body frame? The two frames are shown side by side in *Mathematica* notebook as9-rotations-body-frame.nb. Stare at the animation and imagine that you are moving around with the body and try to get a feeling for how you would see the blue line representing $\vec{\omega}$. Naturally, $\hat{e}_1$ and therefore $\vec{\omega}_1$ doesn't change, because $\hat{e}_1$ in defined to be fixed in the body frame. But $\vec{\omega}_2$, which was fixed in the space frame, is not fixed in the body frame, precisely because of the nonzero $\vec{\omega}_1$, as shown below.



(39)

In the body frame, at $t = 0$, the $\vec{\omega}$ vector looks the same as it did in the space frame. But now at later times, the $\vec{\omega}$ vector rotates around the red $\hat{e}_1$ axis with angular velocity $-\vec{\omega}_1$, so the perpendicular component, $\vec{\omega}_2$ executes uniform circular motion about the red $\hat{e}_1$ axis with angular velocity $-\vec{\omega}_1$, and the rate of change of the vector $\vec{\omega}$ is

$$\frac{\delta}{\delta t}\vec{\omega} = -\vec{\omega}_1 \times \vec{\omega} = -\vec{\omega}_1 \times (\vec{\omega}_1 + \vec{\omega}_2) = -\vec{\omega}_1 \times \vec{\omega}_2 \tag{40}$$

which is the same as (34). Again, you can verify this explicitly. In the body frame

$$\vec{\omega}_1 = \omega_1 \hat{x} \qquad \vec{\omega}_2 = \omega_2 \hat{z} \cos\omega_1 t + \omega_2 \hat{y} \sin\omega_1 t \tag{41}$$

$$\vec{\omega} = \omega_1 \hat{x} + \omega_2 \hat{z} \cos\omega_1 t + \omega_2 \hat{y} \sin\omega_1 t \tag{42}$$

$$\frac{d}{dt}\vec{\omega} = -\omega_1\omega_2 \hat{z} \sin\omega_1 t + \omega_1\omega_2 \hat{y} \cos\omega_1 t = \vec{\omega}_2 \times \vec{\omega}_1 \tag{43}$$

Note that the actual functions of t in (42) (43) look different from those in (40) for $t > 0$, because the coordinate systems only agree for $t = 0$. But they are describing the same **vector relation**, (34).

Before going on, let me just remind you that what $\vec{\omega}$ in the body frame is is **NOT** the angular velocity of the body in a coordinate system fixed in the body. That would of course be zero. Rather $\vec{\omega}$ in the body frame is the true angular velocity of the body in the space frame but described in terms the coordinates of the body frame. When I am talking about the body frame I will try to remember to refer to $\vec{\omega}$ as the $\vec{\omega}$ vector rather than the "angular velocity" to emphasize this important difference.

The free symmetric top in the body frame

I hope that it will not surprise you that our principal tool in analyzing the free symmetric top in the body frame will be the general relation between the change of a vector with respect to the space frame and its change with respect to the body frame, (28)

$$\frac{d\vec{A}}{dt} = \frac{\delta\vec{A}}{\delta t} + \vec{\omega}\times\vec{A}\,.$$

The other thing we need is the vector relation we derived in (13) and (10) between $\vec{L}$, $\vec{\omega}$ and $\hat{e}_3$:

$$\vec{\omega} = \vec{L}/I_\perp - \Omega\hat{e}_3 \tag{13}$$

with Ω given by (10)

$$\Omega = \frac{(I_3 - I_\perp)\,\omega_3}{I_\perp} \tag{10}$$

This vector relation is true in any coordinate system.

We can now calculate how $\vec{\omega}$ and $\vec{L}$ change in the body frame. We will do this in two ways. First we can use (13) and the fact we derived last time that with respect to the space frame the plane of $\vec{\omega}$ and $\hat{e}_3$ is rotating about $\vec{L}$ with angular velocity $\vec{L}/I_\perp$ to write

$$\frac{\delta\vec{\omega}}{\delta t} = \frac{d\vec{\omega}}{dt} = \frac{\vec{L}}{I_\perp}\times\vec{\omega} \tag{44}$$

Again this is a vector equation true in any coordinate system, so we can use it in the space frame, together with (13) to write

$$\frac{\delta\vec{\omega}}{\delta t} = \left(\vec{\omega} + \Omega\hat{e}_3\right)\times\vec{\omega} \tag{45}$$

$$= \vec{\omega}\times\vec{\omega} + \Omega\hat{e}_3\times\vec{\omega} = \Omega\hat{e}_3\times\vec{\omega} \tag{46}$$

But because $\hat{e}_3$ is fixed in the body frame, (46) implies that with respect to the body frame, $\vec{\omega}$ is rotating about $\hat{e}_3$ with angular velocity $\Omega\hat{e}_3$. This is the same trick we used in the space frame, except that now it is $\hat{e}_3$ that is fixed, so we want to think of the rotation as about $\hat{e}_3$. Note that (46) implies that ω_3 is constant just as in the space frame because

$$\frac{\delta}{\delta t}(\hat{e}_3\cdot\vec{\omega}) = \hat{e}_3\cdot\frac{\delta\vec{\omega}}{\delta t} = \hat{e}_3\cdot\left(\Omega\hat{e}_3\times\vec{\omega}\right) = 0\,, \tag{47}$$

so we can independently derive the fact that the coefficients in in the vector equations (13) are fixed.

Alternatively, we can use the fact that $\vec{L}$ is fixed in the space frame and write

$$0 = \frac{\delta\vec{L}}{\delta t} + \vec{\omega}\times\vec{L}\,, \tag{48}$$

so that

$$\frac{\delta\vec{L}}{\delta t} = -\vec{\omega}\times\vec{L} = -\left(\vec{L}/I_\perp - \Omega\hat{e}_3\right)\times\vec{L} = \Omega\hat{e}_3\times\vec{L} \tag{49}$$

This expresses the same physics. With respect to the body frame, $\vec{L}$, like $\vec{\omega}$ must rotate about the $\hat{e}_3$ axis with angular velocity $\Omega\hat{e}_3$ because the entire plane in which $\vec{\omega}$ and $\vec{L}$ are fixed is rotating.

There is a nice way of understanding all this that may help to cement one of the important lessons of the last few weeks — while rotations are complicated and do not add like ordinary vectors, infinitesimal rotations and angular velocities do add. In this case, what we have with respect to the body frame is that the plane formed by the three vectors $\vec{L}$, $\vec{\omega}$ and $\hat{e}_3$ rotates with angular velocity

$$\vec{\omega}_{pb} = \Omega \hat{e}_3 \,. \tag{50}$$

With respect to the space frame, the plane formed by the three vectors $\vec{L}$, $\vec{\omega}$ and $\hat{e}_3$ rotates with angular velocity

$$\vec{\omega}_{ps} = \vec{L}/I_\perp \tag{51}$$

The relation between the space frame and the body frame is that the angular velocity of the plane with respect to the space frame, $\vec{\omega}_{ps}$, is the angular velocity of the plane with respect to the body frame, $\vec{\omega}_{pb}$, plus the angular velocity of the body, $\vec{\omega}$:

$$\vec{\omega}_{ps} = \vec{\omega}_{pb} + \vec{\omega} \tag{52}$$

Indeed, because of (13), this works. Angular velocities add like ordinary vectors.

Euler's equations

There is another approach to the problem of the free symmetric top, which is also more generally useful. We can derive what are called "Euler's equations" by considering (28) where $\vec{A}$ is the angular momentum, $\vec{L}$. We can then use the fact that the torque, $\vec{\tau}$, is the rate of change of angular momentum with respect to the space frame. Using (28), we can write this in terms of the rate of change of $\vec{L}$ with respect to the body frame. Then we can write

$$\vec{\tau} = \frac{d\vec{L}}{dt} = \frac{\delta\vec{L}}{\delta t} + \vec{\omega} \times \vec{L} \tag{53}$$

As we discussed last time, we can evaluate a vector equation in any coordinate system. It is interesting to consider (53) in the body frame in which $\vec{\vec{I}}$ is fixed. Then in the usual basis in which we use the principal axes of our body, $\hat{e}_1$ $\hat{e}_2$ and $\hat{e}_3$, the moment of inertia tensor is

$$\vec{\vec{I}} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{54}$$

and we can express the components of $\vec{L}$ in this basis in terms of the components of $\vec{\omega}$ as

$$L_1 = I_1\, \omega_1 \qquad L_2 = I_2\, \omega_2 \qquad L_3 = I_3\, \omega_3 \tag{55}$$

We now expand the cross product in (53) to get

$$\begin{aligned}\tau_1 &= \frac{\delta L_1}{\delta t} + \omega_2\, L_3 - \omega_3\, L_2 \\ \tau_2 &= \frac{\delta L_2}{\delta t} + \omega_3\, L_1 - \omega_1\, L_3 \\ \tau_3 &= \frac{\delta L_3}{\delta t} + \omega_1\, L_2 - \omega_2\, L_1\end{aligned} \tag{56}$$

Now use (55), $L_1 = I_1\,\omega_1\,,\ L_2 = I_2\,\omega_2\,,\ L_3 = I_3\,\omega_3\,$, to write

$$\begin{aligned}\tau_1 &= I_1\,\frac{\delta\omega_1}{\delta t} + \omega_2\, I_3\,\omega_3 - \omega_3\, I_2\,\omega_2 \\ \tau_2 &= I_2\,\frac{\delta\omega_2}{\delta t} + \omega_3\, I_1\,\omega_1 - \omega_1\, I_3\,\omega_3 \\ \tau_3 &= I_3\,\frac{\delta\omega_3}{\delta t} + \omega_1\, I_2\,\omega_2 - \omega_2\, I_1\,\omega_1\end{aligned} \tag{57}$$

or simplifying slightly

$$\begin{aligned}\tau_1 &= I_1\,\frac{\delta\omega_1}{\delta t} - (I_2 - I_3)\,\omega_2\,\omega_3 \\ \tau_2 &= I_2\,\frac{\delta\omega_2}{\delta t} - (I_3 - I_1)\,\omega_3\,\omega_1 \\ \tau_3 &= I_3\,\frac{\delta\omega_3}{\delta t} - (I_1 - I_2)\,\omega_1\,\omega_2\end{aligned} \tag{58}$$

These are Euler's equations, relating the rate of change of the angular velocity in the body frame to the torque and angular velocity in the body frame and the principal moments of inertia. We might have expected that the right hand side would be proportional to products of two components of $\vec{\omega}$. This ensures that if the angular velocity is in the direction of one of the principal axes, the right hand terms of all the Euler equations vanish because two of the components of $\vec{\omega}$ are zero. And then if the torque vanishes, $\delta\vec{\omega}/\delta t = 0$, because as we know, the body can rotate freely about a principal axis.

The free symmetric top ala Euler

For the free symmetric top, floating out in space and rotating freely about its center of mass with $I_1 = I_2 = I_\perp$, the Euler's equations take the following form:

$$
\begin{aligned}
I_\perp \frac{\delta \omega_1}{\delta t} &= -(I_3 - I_\perp)\, \omega_3\, \omega_2 \\
I_\perp \frac{\delta \omega_2}{\delta t} &= (I_3 - I_\perp)\, \omega_3\, \omega_1 \\
I_3 \frac{\delta \omega_3}{\delta t} &= 0
\end{aligned}
\tag{59}
$$

We immediately conclude that ω_3, the component of angular velocity in the direction of the unequal principal axis, is constant in time relative to the body frame. With ω_3 then fixed, the other two equations take the form

$$
\begin{aligned}
\frac{\delta \omega_1}{\delta t} &= -\frac{(I_3 - I_\perp)\, \omega_3}{I_\perp}\, \omega_2 = -\Omega\, \omega_2 \\
\frac{\delta \omega_2}{\delta t} &= \frac{(I_3 - I_\perp)\, \omega_3}{I_\perp}\, \omega_1 = \Omega\, \omega_1
\end{aligned}
\tag{60}
$$

where

$$\Omega = \frac{(I_3 - I_\perp)\, \omega_3}{I_\perp} \tag{61}$$

and Ω is a constant.

We know that we can rewrite (60) as

$$\frac{\delta \vec{\omega}}{\delta t} = (0, 0, \Omega) \times \vec{\omega} \tag{62}$$

In agreement with the previous analysis in (46), this is what we expect for a vector $\vec{\omega}$ undergoing uniform circular motion with angular velocity Ω about the 3 axis.

You can also see this in a more boring way by solving the coupled differential equations. If we put the second of the two equations in (60) into the first, we get an equation of motion for ω_1,

$$\frac{\delta^2 \omega_1}{\delta t^2} = -\Omega^2\, \omega_1 \tag{63}$$

You can easily check that ω_2 satisfies the same equation. (63) is just the equation for simple harmonic motion, with the general solution

$$\omega_1(t) = a \cos(\Omega t + \phi) \tag{64}$$

Then (60) implies

$$\omega_2(t) = a \sin(\Omega t + \phi) \tag{65}$$

As usual, a and ϕ must be set by the initial conditions. So, as we saw directly from (62), (64) and (65) describe the two-dimensional vector, $\big(\omega_1(t), \omega_2(t)\big)$ undergoing uniform circular motion with angular velocity Ω. Note that the sign of Ω is related to the sign of ω_3, and also whether $I_\perp > I_3$ (which is what we would get for a prolate spheroid) or $I_\perp < I_3$ (for an oblate spheroid like the earth).

Thus the angular velocity in the body frame is fixed in magnitude, but changes direction, moving with constant angular velocity around a cone.

The tennis racket theorem

Euler's equations can also be used to understand an interesting fact about rotating rigid bodies with no two equal moments of inertia. I'm a tennis player, so the tennis racket theorem is a favorite of mine. The tennis racket, in this case, is just a convenient example of a rigid body with three unequal moments of inertia. We know that a tennis racket, like any rigid body, is happy to rotate freely about any one of its principal axes. For such a rotation, the angular momentum and the angular velocity vector are in the same direction. The theorem concerns the nature of free rotations (no torques) that begin with an angular velocity almost, but not quite, along one of the principal axes. The statement of the theorem is that if the axis has the largest or the smallest moment of inertia, the motion is stable and the angular velocity precesses about the axis, never getting very far away. However, if the axis has the intermediate moment of inertia, the motion is unstable. The angular velocity moves away from the axis exponentially and the motion becomes rather complicated. Here is how it goes. Suppose that we consider motion in which $\vec{\omega}$ in the body frame is nearly along principal axis 3, so that in our nice body frame, ω_1 and ω_2 are much smaller than ω_3. Euler's equations look like

$$\begin{aligned} I_1 \frac{\delta\omega_1}{\delta t} &= (I_2 - I_3)\,\omega_2\,\omega_3 \\ I_2 \frac{\delta\omega_2}{\delta t} &= (I_3 - I_1)\,\omega_3\,\omega_1 \\ I_3 \frac{\delta\omega_3}{\delta t} &= (I_1 - I_2)\,\omega_1\,\omega_2 \end{aligned} \tag{66}$$

For the assumed initial conditions, we can generally ignore the time dependence of ω_3, because its time derivative is proportional to the product of two small numbers, ω_1 and ω_2. This argument could go wrong if I_3 is very close to either I_1 or I_2, but for a tennis racket in which the three principal moments are very different, it is a good approximation.

If we then treat ω_3 as a constant, the other two Euler equations are linear in the small quantities,

ω_1 and ω_2, and we can write them (approximately) as.

$$\begin{aligned}\frac{\delta\omega_1}{\delta t} &= \frac{1}{I_1}\,(I_2 - I_3)\,\omega_3\,\omega_2 \\ \frac{\delta\omega_2}{\delta t} &= \frac{1}{I_2}\,(I_3 - I_1)\,\omega_3\,\omega_1\end{aligned} \tag{67}$$

Differentiate the first of these and use the second to get

$$\frac{\delta^2\omega_1}{\delta t^2} = \frac{1}{I_1}\,(I_2 - I_3)\,\omega_3\,\frac{\delta\omega_2}{\delta t} = \frac{1}{I_1 I_2}\,(I_3 - I_1)\,(I_2 - I_3)\,\omega_3^2\,\omega_1 \tag{68}$$

Now here is the point. If I_3 is the largest or smallest moment, then

$$\frac{1}{I_1 I_2}\,(I_3 - I_1)\,(I_2 - I_3)\,\omega_3^2 < 0 \tag{69}$$

But then (68) describes an oscillation. ω_1 oscillates about zero and remains small. The same thing happens to ω_2. In this case, the motion of the system remains basically about axis 3 with small oscillating wobbles. However if I_3 is the intermediate moment, then

$$\frac{1}{I_1 I_2}\,(I_3 - I_1)\,(I_2 - I_3)\,\omega_3^2 > 0 \tag{70}$$

Then (68) describes an unstable equilibrium and ω_1 and ω_2 grow exponentially. Then our approximation quickly breaks down. We can't tell what happens next without a more detailed analysis. But what is clear is that the motion around the axis with the intermediate moment of inertia is much more complicated. You can easily see this yourself by playing with a tennis racket, or any similarly shaped object.

lecture 22

Topics:
- Where are we now?
- Derivation of fictitious forces
- Precession of tops
- Nutation
- Tidal forces
- Extra Dimensions

Where are we now?

This week, the general subject is gravity (and more generally "central" forces) and "fictitious forces." Fictitious forces are convenient constructions that allow us to continue to use $\vec{F} = m\vec{a}$ in non-inertial frames of reference. This is actually familiar to all of us. Our brain makes the same kind of construction whenever we turn a corner in a moving car and we feel a centrifugal force. Gravity is also familiar, but seems rather different. Einstein however, would tell us that the difference between the two is not so obvious as it seems. It is not an accident, he would say, that both forces are proportional to the mass of the particle on which they act. We will begin to see why this is important when we talk about the tides. And it will come back again later when we talk about cosmology.

We will probably not get to everthing in the notes in the lecture today, but some of it is better to just read about. I will try to take a little time at the end to talk a bit about the subject of "extra dimensions" and the connection with gravity, because this crazy idea is sort of fun and something that many people are interested in today.

Derivation of fictitious forces

The crucial input to the study of fictitious forces is that the motion of an arbitrary accelerating frame, like an arbitrarily moving rigid body, can be described at any time by the motion of its origin, $\dot{\vec{R}}$, and by its angular velocity of rotation about the origin, $\vec{\omega}$. Many of the same ideas that go into thinking about the physics of rotating rigid bodies can be taken over directly to the study of rotating reference frames. The frame of reference in which the rigid body is at rest is obviously a good one to think about when discussing a rigid body, and if the body is rotating, then of course this frame is rotating too.

The basic idea of fictitious forces is this. If I insist on using an accelerating and/or rotating reference frame, then particles that are moving at constant velocity in free space will appear to be accelerating. In order to make $\vec{F} = m\vec{a}$ work in such an non-inertial frame, we simply interpret this acceleration as being due to a force, called the fictitious force. More specifically, we break $m\,a$ up into a piece associated with the frame and all the rest and move the "frame" part to the other side of the equation, treating it as a force:

$$\vec{F}_{\substack{\text{on body}\\ \text{wrto}\\ \text{inertial}\\ \text{frame}}} = m\,\vec{a}_{\substack{\text{of body}\\ \text{wrto}\\ \text{inertial}\\ \text{frame}}} = m\,\vec{a}_{\substack{\text{of body}\\ \text{wrto}\\ \text{accel.}\\ \text{frame}}} + m\,\vec{a}_{\substack{\text{accel.}\\ \text{frame}}} \tag{1}$$

$$\text{“}\vec{F}_{\substack{\text{on body}\\ \text{wrto}\\ \text{accel.}\\ \text{frame}}}\text{”} = \vec{F}_{\substack{\text{on body}\\ \text{wrto}\\ \text{inertial}\\ \text{frame}}} \underbrace{-m\,\vec{a}_{\substack{\text{accel.}\\ \text{frame}}}}_{\substack{\text{fictitious}\\ \text{force}}} = m\,\vec{a}_{\substack{\text{of body}\\ \text{wrto}\\ \text{accel.}\\ \text{frame}}} \tag{2}$$

The fictitious force is always proportional to the mass of the particle, because of the m in $\vec{F} = m\,\vec{a}$.

There were two basic kinds of questions about this on the QA. Do we really have all the fictional forces? AND Why do we call them "fictional"? These are closely related, it turns out. We call these forces "fictional" because they can be eliminated completely everywhere by going to an appropriate reference frame. If they can be completely elimnated in this way, then in some sense they are just due to our poor choice of reference, and not to any "real" physics. But this is a little tricky. The earth's gravity is definitely "real" even though we don't feel it if we are in a space ship in orbit above the earth. What is the difference between a fictitious force and gravity? Both of them give a force on a point particle proportional to the particle's mass. This is why we don't feel gravity in orbit. It's effect is canceled by the fictitous translational force. In fact, the only way we can tell the difference is by looking at that form of the force **everywhere.** When we define a coordinate system, we are laying down a grid that allows us (theoretically - we can't really do this of course) to find the coordinates of any point anywhere in space (remember that we are doing Newtonian mechanics and ignoring general relativity). Then we define our frame by specifying the coordinate systen at each time. Each coordinate system is a rigid thing — not physically, but theoretically. So like a rigid body, it is specified by only six parameters, the three coordinates of the origin, and the three angles that determine the orientation of the coordinate axes. This is why we know that we have all the fictitious forces. We know from our study of rigid bodies that the most general motion of the rigid coordinate system is equivalent to a translation and a rotation, so we know the form of our fictitious forces **everywhere** and we can eliminate them **everywhere** by going to an appropriate frame. Gravity doesn't work this way. We can eliminate gravity at any point by going to an accelerating coordinate frame in which our coordinate system is "falling" along with us. But this cancellation of gravity only works at one point. That is why we say that gravity is real.

Relativity (as usual) makes things even more complicated. Once we realize that gravity curves space the whole notion of coordinate system is more complicated. But as long as we say away from black holes and that sort of thing, we can coninue to use the Newtonian picture at least approximately, and that is what we will do.

Here is a brief derivation of the fictitious forces. It is equivalent to what you will read about in Morin's book, but it breaks the transformation to the accelerating coordinate system up into two steps, for me at least, this makes it a little simpler to understand. For a coordinate system that is rotating with angular velocity $\vec{\omega}$ about its origin and whose origin is translated by a vector $\vec{R}$ (where $\vec{\omega}$ and $\vec{R}$ have arbitrary time dependence) with respect to some fixed point in an inertial frame, the fictitious force is the sum of four terms:

$$\begin{array}{ll} \textit{Centrifugal} & -m\,\vec{\omega}\times(\vec{\omega}\times\vec{r}\,) \\ \textit{Coriolis} & -2m\,\vec{\omega}\times\vec{v} \\ \textit{Azimuthal} & -m\,\dot{\vec{\omega}}\times\vec{r} \\ \textit{Translational} & -m\,\dfrac{d^2\vec{R}}{dt^2} \end{array} \tag{3}$$

As I said, I think it is easiest to demonstrate (3) in two steps. To go from a coordinate system fixed in space to one undergoing arbitrary acceleration and rotation, first go from the fixed coordinate system to a translating coordinate system that is not rotating with respect to the fixed system, but with its origin accelerating along with the moving coordinate system. We want to know what happens to $\vec{F}=m\vec{a}$ when we do this. Let's call the original space frame force and coordinates $\vec{F}_s$ and $\vec{r}_s$, and the translating frame force and coordinates $\vec{F}_t$ and $\vec{r}$, so that we begin with

$$\vec{F}_s = m\,\frac{d^2}{dt^2}\,\vec{r}_s \tag{4}$$

Then the new coordinates are

$$\vec{r} = \vec{r}_s - \vec{R} \tag{5}$$

and we can write

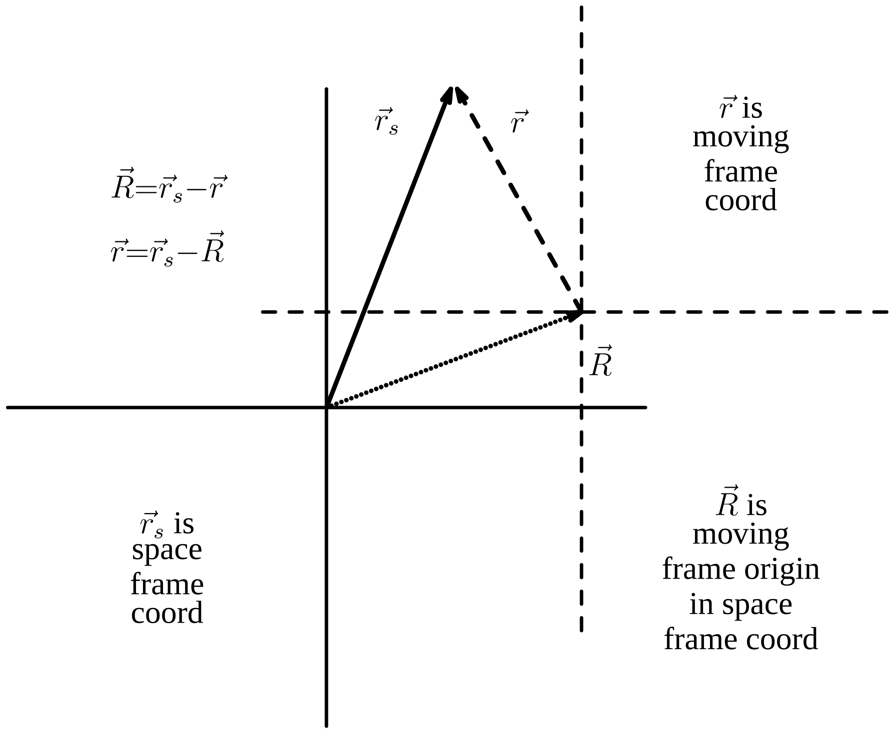
$$m\,\frac{d^2}{dt^2}\vec{r} = m\,\frac{d^2}{dt^2}\left(\vec{r}_s - \vec{R}\right) = \vec{F}_s - m\,\frac{d^2}{dt^2}\vec{r}_s \tag{6}$$

So that

$$\vec{F}_t \equiv \vec{F}_s - m\,\frac{d^2}{dt^2}\,\vec{R} = m\,\frac{d^2}{dt^2}\,\vec{r} \tag{7}$$

Thus the equation of motion in the translating coordinate system is the same as in the space frame, except that the force has a new piece – the fictitious translational force.

It may help to see the relation (5) represented pictorially. In the diagram below the solid axes and arrow represent the coordinates of a point in the space frame, $\vec{r}_s$. The dashed axes and arrow represent the coordinates of a point in the translating frame, $\vec{r}$. The dotted arrow is $\vec{R}$, which as you can see describes the coordinates of the origin of the translating frame in the space frame.



(8)

We can now do the rest of the job, and go from a coordinate system (that we will call the translating frame) to another with the same origin, but that is rotating arbitrarily. We will continue to call

$$\frac{d}{dt} \tag{9}$$

the derivative that describes how vectors change with respect to translating frame and will indicate the rate of change of a vector in the arbitrary frame by $\delta/\delta t$ as we did in our study of rigid bodies. Now we can use the relation that is familiar from our study of rigid bodies,

$$\frac{d\vec{A}}{dt} = \frac{\delta\vec{A}}{\delta t} + \vec{\omega}\times\vec{A}\,. \tag{10}$$

Let $\vec{A} = \vec{r}$ and (10) becomes

$$\frac{d\vec{r}}{dt} = \frac{\delta\vec{r}}{\delta t} + \vec{\omega}\times\vec{r}\,. \tag{11}$$

Now we can differentiate again in the translating frame and repeatedly make use of (11) —

$$\frac{d^2\vec{r}}{dt^2} = \frac{d}{dt}\frac{\delta\vec{r}}{\delta t} + \frac{d}{dt}(\vec{\omega}\times\vec{r})\,. \tag{12}$$

$$= \frac{\delta^2\vec{r}}{\delta t^2} + \vec{\omega}\times\frac{\delta\vec{r}}{\delta t} + \left(\frac{d}{dt}\vec{\omega}\right)\times\vec{r} + \vec{\omega}\times\frac{d}{dt}\vec{r}. \tag{13}$$

$$= \frac{\delta^2\vec{r}}{\delta t^2} + \vec{\omega}\times\frac{\delta\vec{r}}{\delta t} + \frac{\delta\vec{\omega}}{\delta t}\times\vec{r} + \vec{\omega}\times\frac{\delta\vec{r}}{\delta t} + \vec{\omega}\times(\vec{\omega}\times\vec{r})\,. \tag{14}$$

so finally

$$\frac{d^2\vec{r}}{dt^2} = \frac{\delta^2\vec{r}}{\delta t^2} + 2\vec{\omega}\times\vec{v} + \frac{\delta\vec{\omega}}{\delta t}\times\vec{r} + \vec{\omega}\times(\vec{\omega}\times\vec{r}) \ . \tag{15}$$

Now we will replace the $\delta/\delta t$s by dots just to make it look simpler.

$$\frac{d^2\vec{r}}{dt^2} = \ddot{\vec{r}} + 2\vec{\omega}\times\vec{v} + \dot{\vec{\omega}}\times\vec{r} + \vec{\omega}\times(\vec{\omega}\times\vec{r}) \ . \tag{16}$$

where we have used the fact that $\vec{v} = \dot{\vec{r}}$ – the velocity with respect to the arbitrary frame.

Finally we just have to put (16) into (7) and move the "fictitious" terms to the other side to get the equation of motion in the arbitrary frame,

$$\vec{F}_a = m\,\ddot{\vec{r}} \tag{17}$$

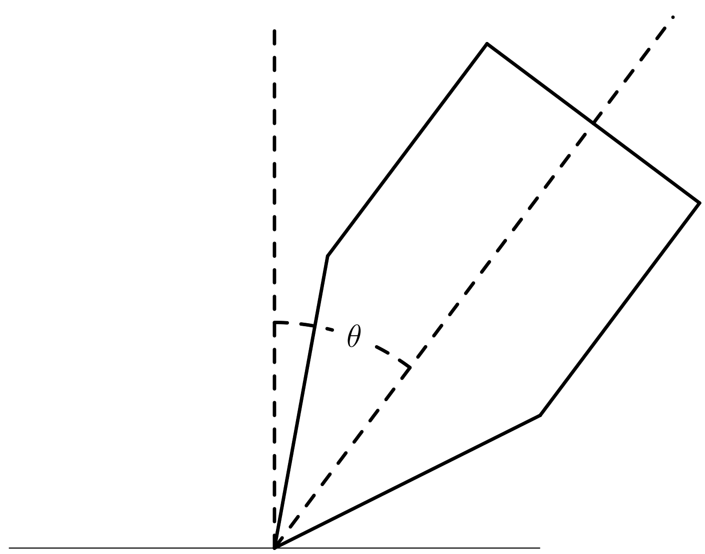
where

$$\vec{F}_a = \vec{F}_s - m\,\vec{\omega}\times(\vec{\omega}\times\vec{r}) - 2m\,\omega\times\vec{v} - m\,\dot{\vec{\omega}}\times\vec{r} - m\,\frac{d^2}{dt^2}\,\vec{R}\,, \tag{18}$$

which is what we mean by (3). We we spend a lot of time next week discussing the physics of each of the terms in (18).

Precession of tops

Let's first see how precession works in quantitative detail. The system is illustrated below:

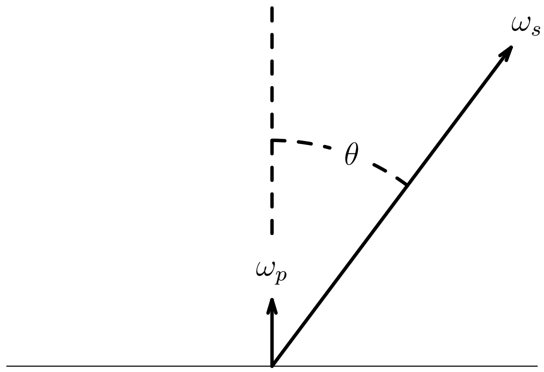


(19)

Suppose that at some angle θ, the torque is τ — into the paper in the figure. We will actually ignore the dependence of the torque on θ and simply assume that the magnitude of the torque is constant and equal to τ. This will allow us to understand what is happening more easily, and it does not change the qualitative nature of the physics, just some boring details.

Now we will first consider a motion in which the top is started with exactly the right precessional velocity to keep it at the angle θ. In this case, the angular velocity has two contributions.

There is a large contribution, ω_s, in the direction of the symmetry axis of the top — because the top is rapidly spinning. There is a vertical contribution, ω_p, due to the precession caused by the torque. Typically, for a rapidly rotating top, the torque is small in the sense that the precession frequency ω_p is much smaller than the angular velocity ω_s associated with the spin of the top.



(20)

From (20), we see that the component of angular momentum in the direction of the symmetry axis is

$$L_{\parallel} = I_3\,\omega_3 \tag{21}$$

where ω_3 is the total component of angular velocity along the symmetry axis.

$$\omega_3 = \omega_s + \omega_p \cos\theta \tag{22}$$

The quantity ω_3 is actually somewhat more interesting than ω_s because it is conserved if the tip of the top is frictionless. The component of angular momentum perpendicular to the symmetry axis is

$$L_{\perp} = I_{\perp}\,\omega_p \sin\theta \tag{23}$$

What we care about for the precession is the horizontal component of angular momentum, which is

$$L_x = \sin\theta\,L_{\parallel} - \cos\theta\,L_{\perp} = \sin\theta\,I_3\,\omega_3 - \sin\theta\cos\theta\,I_{\perp}\,\omega_p \tag{24}$$

Normally, we ignore the small second term on the right hand side of (24), and do not distinguish between ω_s and ω_3. But I just wanted to show you that it is possible to take into account the contribution from the precession if we want to. At any rate, the rate of change of angular momentum for uniform precessional motion is $\omega_p\,L_x$, and thus the condition for uniform precession is

$$\tau = \omega_p\,L_x = \omega_p\left(\sin\theta\,I_3\,\omega_3 - \sin\theta\cos\theta\,I_{\perp}\,\omega_p\right) \tag{25}$$

This is a quadratic equation for ω_p with solutions

$$\omega_p = \frac{I_3\,\omega_3}{2I_{\perp}\cos\theta} \pm \sqrt{\left(\frac{I_3\,\omega_3}{2I_{\perp}\cos\theta}\right)^2 - \frac{\tau}{\sin\theta\cos\theta\,I_{\perp}}} \tag{26}$$

The plus sign here gives what is called fast precession, which is precession of the order of the rotational velocity of the top. If the torque is small, we can write it as

$$\omega_p = \frac{I_3\,\omega_3}{2I_\perp\cos\theta} + \sqrt{\left(\frac{I_3\,\omega_3}{2I_\perp\cos\theta}\right)^2 - \frac{\tau}{\sin\theta\cos\theta\,I_\perp}} \approx \frac{I_3\,\omega_3}{I_\perp\cos\theta} \tag{27}$$

This is just a slightly perturbed version of the free rotation of a symmetric top that we have just understood in detail — the torque plays very little role in it. This is not what we usually see, because the top usually doesn't get started moving that fast. Instead, we are interested in slow precession, the minus sign solution, which gives

$$\begin{aligned}
\omega_p &= \frac{I_3\,\omega_3}{2I_\perp\cos\theta} - \sqrt{\left(\frac{I_3\,\omega_3}{2I_\perp\cos\theta}\right)^2 - \frac{\tau}{\sin\theta\cos\theta\,I_\perp}} \\
&= \frac{I_3\,\omega_3}{2I_\perp\cos\theta}\left(1 - \sqrt{1 - \frac{4I_\perp\tau\cos\theta}{\sin\theta\,\omega_3^2 I_3^2}}\right) \\
&\approx \frac{I_3\,\omega_3}{2I_\perp\cos\theta}\left(1 - 1 + \frac{2I_\perp\tau\cos\theta}{\sin\theta\,\omega_3^2 I_3^2}\right) = \frac{\tau}{\sin\theta\,I_3\,\omega_3}
\end{aligned} \tag{28}$$

This is the usual result in which the precession frequency is proportional to the torque and inversely proportional to the angular momentum.

Nutation

So if the top is moving at exactly the right angular velocity, it precesses at constant angular velocity and stays at the same angle. But what happens if things are not quite right? In particular, suppose that we are holding the top at some fixed angle and suddenly drop it. This is a typical situation. What happens?

The best way to understand what happens is to go to an accelerating frame that is rotating around the pivot point of the top with angular velocity ω_p. This is not an inertial frame, of course, which means that there are so-called fictitious forces that make up for the fact that the frame is accelerating. We will talk about these in detail later. At any rate, in this frame, the motion with constant θ and precession rate ω_p just looks like a top with its symmetry axis sitting still in space. This means that in this frame there must be a fictitious force that exactly cancels the torque, so that the angular momentum is conserved. I presume that this a basically a coriolis force, but it doesn't matter what it is because it must be there.

Now what does the process look like in this rotating frame if we drop the top from rest in the space frame? In the moving frame, because the symmetry axis of the top was initially at rest in the space frame, it is initially rotating with angular velocity $-\omega_p$ in the rotating frame. But that means that there is a small additional component to the angular velocity in the vertical direction, and thus the angular velocity and the angular momentum of the top are slightly displaced from the symmetry axis. Because there is no torque in this special frame, this just reduces the free rotation problem that we have already analyzed. The symmetry axis of the top precesses rapidly around the angular momentum with frequency $L/I_\perp$. Back in the space frame, this motion is superimposed on the generally much slower precession of the angular momentum produced by the torque. The rapid motion is called nutation.

Not only is the nutational motion very rapid, but the amplitude is usually very small. In the case we discussed in which the top is dropped from rest, the angle of the angular velocity in the special frame to the symmetry axis is of order ω_p/ω_3. For a rapidly rotating top, the motion is often to small and too rapid to see. But you can feel it or hear it under the right circumstances.

As I warned you I would, I have ignored the dependence of the torque on θ in discussing nutation. I hope that you can now see why this doesn't make much difference. The torque is a very small effect that produces the slow precession of the plane of $\hat{e}_3$. The nutation results if the angular momentum is not quite lined up with the symmetry axis. These two effects have very little to do with one another. Small changes in the torque as $\hat{e}_3$ precesses rapidly about $\vec{L}$ will have a very small effect on the motion.

Tidal forces

The phenomenon of tidal forces is rather special to gravitation. It is also both physically important and a nice example of the Taylor expansion with several variables, so I can't resist talking about it (though I might skip it or put it off til next week if we get behind). The idea is to ask what the gravity of a distant object does to a mass on the surface of an approximately spherical planet like the earth (here we ignore the oblateness, which has only a tiny effect on this physics). Suppose the center of the sphere is at a point $\vec{r}_0$. Now by Newton's theorem, we know that the gravitational force from a distant body on a mass m at $\vec{r}$ is given by

$$-G\,M\,m\,\frac{\vec{r}-\vec{\rho}}{|\vec{r}-\vec{\rho}|^3} \tag{29}$$

where $\vec{\rho}$ and M are the position and mass of the distant body, so you might think that all we have to do is to evaluate (29) for $\vec{r}$ on the surface of a sphere, given by $|\vec{r}-\vec{r}_0| = R$, where R is the radius of the sphere. But that is not quite what we are interested in for the gravitational force. The gravitational force acts not only on the mass m, but also on the planet, where the forces acts on the center. What we are actually interested in is the accelerated coordinate system that is moving along with the planet, with the center $\vec{r}_0$ fixed. Thus we must add to (29), the fictitious translational force associated with the acceleration of the planet due to gravity. The acceleration is

$$-G\,M\,\frac{\vec{r}_0-\vec{\rho}}{|\vec{r}_0-\vec{\rho}|^3} \tag{30}$$

and the fictitious translational force is

$$G\,M\,m\,\frac{\vec{r}_0-\vec{\rho}}{|\vec{r}_0-\vec{\rho}|^3} \tag{31}$$

The interesting thing is that this is trying to cancel the effect of (29) and we are only interested in the difference, which is very small.

$$\vec{F}_{\text{tidal}} = -G\,M\,m\,\frac{\vec{r}-\vec{\rho}}{|\vec{r}-\vec{\rho}|^3} + G\,M\,m\,\frac{\vec{r}_0-\vec{\rho}}{|\vec{r}_0-\vec{\rho}|^3} \tag{32}$$

Since in this accelerating coordinate system, $\vec{r}_0$ is fixed, we might as well take it to be the origin of our coordinate system, so we take $\vec{r}_0 = 0$. This makes the formulas simpler. For this more convenient choice, (32) becomes (moving some signs around to make things look simpler)

$$\vec{F}_{\text{tidal}} = G\,M\,m\,\frac{\vec{\rho}-\vec{r}}{|\vec{\rho}-\vec{r}|^3} - G\,M\,m\,\frac{\vec{\rho}}{|\vec{\rho}|^3} \tag{33}$$

Now we can use the fact that $|\vec{r}| \ll |\vec{\rho}|$ and get an even simpler expression by Taylor expanding the first term in powers of $\vec{r}$. This is easier if we first have some fun with vector calculus. We can write

$$\vec{F}_{\text{tidal}} = -G\,M\,m\,\vec{\nabla}_\rho\left(\frac{1}{|\vec{\rho}-\vec{r}|} - \frac{1}{|\vec{\rho}|}\right) \tag{34}$$

Where $\vec{\nabla}_\rho$ is the gradient operator for the vector variable $\vec{\rho}$,

$$\vec{\nabla}_\rho = \left(\frac{\partial}{\partial\rho_x}, \frac{\partial}{\partial\rho_y}, \frac{\partial}{\partial\rho_z}\right) \tag{35}$$

Equation (34) may look scary, but actually it is the usual connection between a $1/r^2$ force and a $1/r$ potential. We get the force by taking minus the gradient of the potential. In this case, we are interested the difference between two forces - the real one from gravity and the fictitious one from the earth's acceleration, and that is minus the gradient of the difference between two potential. In fact, this is a nice example of how a little bit of knowledge of vector calculus can simplify your life a lot. If you had to apply the Taylor expansion separately to each component of (33), it would be at least three times as much work and the result would be more cumbersome to write down.

Now we can Taylor expand the term in parentheses in (34), before we take the gradient - and this is a lot easier

$$\left(\frac{1}{|\vec{\rho}-\vec{r}|} - \frac{1}{|\vec{\rho}|}\right) = \left(\left(\rho^2 - 2\vec{r}\cdot\vec{\rho} + r^2\right)^{-1/2} - \frac{1}{|\vec{\rho}|}\right) = \left(\frac{1}{|\vec{\rho}|} + \frac{\vec{r}\cdot\vec{\rho}}{|\vec{\rho}|^3} + \cdots - \frac{1}{|\vec{\rho}|}\right) \tag{36}$$

The lowest order term in the Taylor expansion of (36) cancels. That is the point — a constant gravitational force does nothing if one is falling along with it. The first order term is

$$\left(\frac{1}{|\vec{\rho}-\vec{r}|} - \frac{1}{|\vec{\rho}|}\right) \approx \frac{\vec{r}\cdot\vec{\rho}}{|\vec{\rho}|^3} \tag{37}$$

Putting this back into (34) gives

$$\vec{F}_{\text{tidal}} = -G\,M\,m\,\vec{\nabla}_\rho\left[\frac{\vec{r}\cdot\vec{\rho}}{|\vec{\rho}|^3}\right] \tag{38}$$

Acting on the $\vec{r}$ in the numerator, the gradient operator in (38) gives

$$-\,G\,M\,m\left[\frac{\vec{r}}{|\vec{\rho}|^3}\right] \tag{39}$$

while acting on the denominator it gives

$$G\,M\,m\left[\frac{3\vec{\rho}\,(\vec{r}\cdot\vec{\rho}\,)}{|\vec{\rho}\,|^5}\right] \tag{40}$$

Thus

$$\vec{F}_{\rm tidal} = \frac{G\,M\,m}{\rho^5}\left[3(\vec{r}\cdot\vec{\rho}\,)\,\vec{\rho}-\rho^2\,\vec{r}\,\right] = \frac{G\,M\,m}{\rho^3}\left[3(\vec{r}\cdot\hat{\rho})\,\hat{\rho}-\vec{r}\,\right] \tag{41}$$

What does this mean? Remember that we are using a coordinate system with the center of the earth as the origin. Thus $\vec{r}$ is a vector from the center of the earth — that is it is straight up out of the ground. And $\vec{\rho}$ is a vector from the earth to the body producing the tidal force. Notice that the sign $\vec{\rho}$ in (41) doesn't matter. This slightly unintuitive result is one of the more amusing features of the tidal force. It is the same on the side closest to the body producing it as on the other side. We will come back to this and try to understand it qualitatively in a moment.

If $\vec{r}$ and $\vec{\rho}$ are parallel, the force is in the positive $\hat{r}$ direction — which means outward. If $\vec{r}$ and $\vec{\rho}$ are perpendicular, it is in the negative $\hat{r}$ — which means inward. Elsewhere, it is not parallel to $\hat{r}$. For

$$3(\vec{r}\cdot\vec{\rho})^2 = \rho^2\,r^2 \tag{42}$$

when the cosine of the angle between $\vec{r}$ and $\vec{\rho}$ is $1/\sqrt{3}$, it is perpendicular to $\hat{r}$ (the angle is about $55°$).

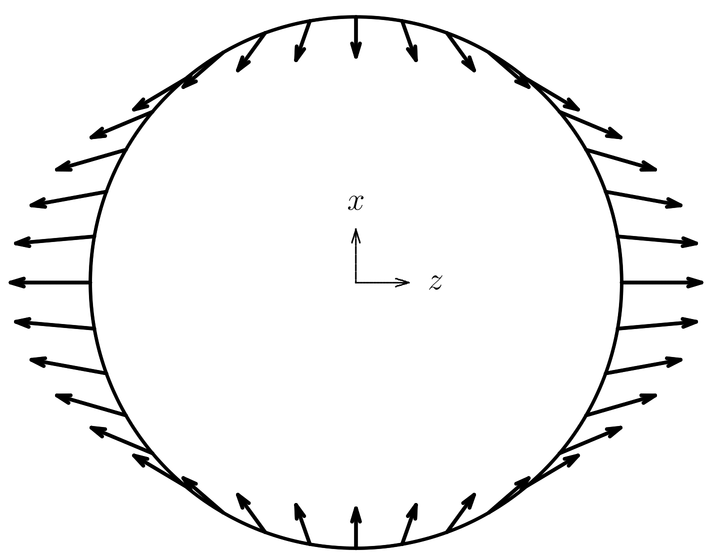
To get a more quantitative description, let us take $\vec{\rho}$ in the $\hat{z}$ direction and look in the x-z plane through the center of the earth. Then we can write

$$\hat{r} = \cos\theta\,\hat{z} + \sin\theta\,\hat{x} \tag{43}$$

and

$$\vec{F}_{\rm tidal} \propto 3(\hat{\rho}\cdot\hat{r})\,\hat{\rho} - \hat{r} = 3(\hat{z}\cdot\hat{r})\,\hat{z} - \hat{r} = 2\cos\theta\,\hat{z} - \sin\theta\,\hat{x} \tag{44}$$

It looks something like this (this is a slice — I have tried to show what it looks like in 3D in the *Mathematica* file **tides.nb**, but I am not sure this has been a real success):



(45)

You can see how this produces the tides, pulling the ocean towards and directly away from the sun or moon.

Let's see if we can understand in our bones why it is that the tidal force is outward, away from the center of the earth, both on the side of the earth closest to the sun, and on the side farthest away.[1] The point is that there are two effects — the gravitational force of the sun, and the "fictitious force", that is the effect of the accelerating coordinate system. At the center of the earth, the two effects cancel, the $1/r^2$ nature of the gravitational force implies that on the side close to the sun, the gravitational attraction wins and the net tidal force is towards the sun, while on the other side, because the gravitational force is weaker, the fictitious force wins, and the tidal force is away from the sun.

This rather odd looking tidal force is characteristic of gravity in free space. It depends crucially on the remarkable feature of gravity that the same quantity — mass — appears in the strength of the force and in the inertial term in Newton's law. It is this that causes us to subtract the force on the center of the earth, to produce what we see in (45). It was this bit of magic that gave Einstein the clue to construct gravity as a geometric theory.

Extra Dimensions

Extra dimensions are fun. We probably won't have too much time to discuss this in lecture today, so I thought I would refer you to an amusing journalistic description of what is called the brane world picture.[2] A braneworld is a modern version of flatland. There are two ideas. The first is this. Suppose that there are space dimensions beyond the usual three. It doesn't make any sense that they are just like the usual three because we would know about them, so suppose that while these extra dimensions are there, we are stuck on the usual three. All the matter that we are made of, and the electric (and strong and weak) forces that make us work the way we do, everything is confined to the usual three dimensions. Now you should object that we have done nothing at all by positing the existence of the extra dimensions, and you would be right. Such extra dimensions would be uninteresting, because we would never see them. But let's try to make something interesting about the fact that we have mentioned before that gravity is much weaker than the other forces. Perhaps matter and the non-gravitational forces - electromagnetism, the strong force and so on, are confined to the usual three dimensions, but gravity gets out into the extra dimensions. This is potentially interesting.

What we mean precisely by saying that gravity gets out into the extra dimensions is that the field lines of gravity are not confined to our three dimensional space, but spread out into the extra dimensions as well.

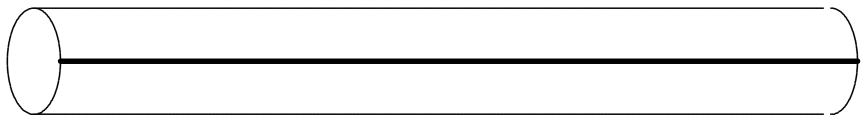
This is the basic picture of a "brane-world." The term "brane" is sort of a generalization of "membrane" and it refers to the usual three space dimensions living in a higher dimensional space. Gravity lives in the full higher dimensional space because it is somehow related to the geometry (this is a bit vague, I know).

[1]The argument is a little easier to give for the sun, so we will do that, though the moon works the same way.

[2]If the link doesn't work for you, I have included the file on the web page as braneworlds.pdf.

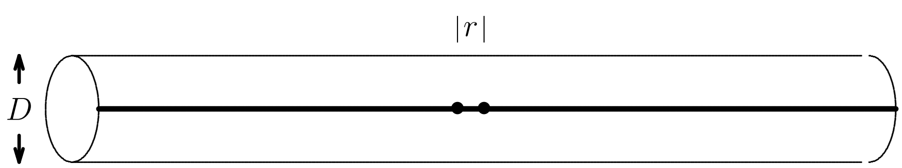
But now you should object for another reason. If the field lines of gravity spread out into the extra dimensions, you would expect the force to fall off faster than $1/r^2$ as the distance increases. But we know that gravity is a $1/r^2$ force, don't we?

This is where the second part of the brane-world idea comes in. The extra dimensions are curled up. A useful picture of this space is a soda straw with a line on it, as shown below

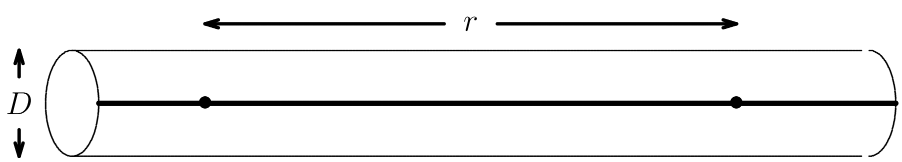


The thick black line is intended to represent the brane on which we live - that is ordinary three dimensional space. Of course, the line is only one dimensional, but if we had more dimensions to draw in, we could include the other two. The important thing is that these dimensions are flat and infinite. We have only shown a segment of the infinitely long straw. The other dimension - around the straw - represents the extra dimensions. Again, there might be more of them. But the important thing is that these extra dimensions are curled up and finite. You have to remember when you stare at this that we have shown just one real and one extra dimension, to make it easy to visualize. But otherwise, it is a really helpful picture.

Now suppose we put two masses on the brane - representing two masses in our three dimensional world. The situation is very different depending on whether the distance r between the masses is large or small compared to D, the diameter of the extra dimensions.



For $r \ll D$, the field lines spread out into the extra dimensions. On this small scale, the multi-dimensional surface of the soda straw looks approximately flat, and gravity looks n dimensional. The force falls off like $1/r^{n-1}$.

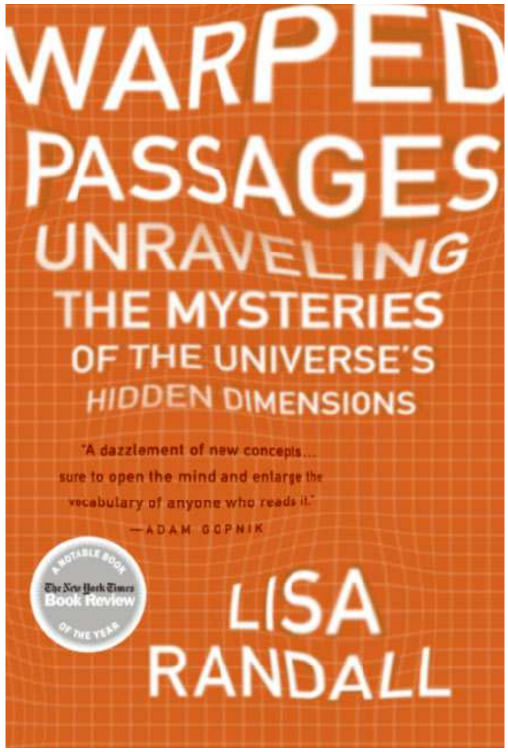


field lines uniformly spread
over extra dimensions

For $r \gg D$, the field lines spread out uniformly over extra the extra dimensions once you get a distance much greater than D from the masses. Then the force looks three dimensional, proportional to $1/r^2$.

One of the reasons that this may be interesting is that it provides an explanation for the weakness of gravity. Gravity is week for masses on our brane world not because it is weak at short distances, but because it is diluted by being spread over the "large" extra dimensions. What "large" might mean is anybody's guess. But it could certainly show up at the LHC if we are lucky.

If this seems interesting to you may want to check out:



lecture 24

Topics:
- Disclaimer
- The Cosmological Principle and the Hubble expansion
- Newton's theorem and the expansion of the Universe
- Hubble dynamics
- Dynamics of a flat universe
- Newton's theorem and dark matter

Disclaimer

I want spend a few lectures on cosmology. This is not an area that I have worked in, so I do not know it in my bones the way I know particle physics. I will do my best to get across the big ideas, but we won't have to go very deeply into things before my understanding starts to get spotty. I'll try to remember to warn you when this is happening. I am certainly in no position to give the latest values of interesting quantities like age of the universe or the Hubble constant from a position of deep knowledge. When I do quote such a number, I will simply take it from the web page of an old friend of mine, Ned Wright, an astrophysicist at UCLA and a leader on the COBE Collaboration. Ned Wright's Cosmology page is a fun site, and I recommend for poking around. What I am going to focus on is things that are directly connected to the stuff we have been talking about in this course.

For the most part, we are going to focus of things that we can say without getting very relativistic. The main reason for this is that the discussion can be much simpler when we can avoid confusing issues like relativity of simultanaity. To do things right would require general relativity (which we will not discuss, though we will take one result when we need it). And it would take so much time to get all the definitions straight that we wouldn't have time left for the fun physics.

At the end of the last lecture on cosmology, which will be the last lecture in the course, I will return to the question that we talked about at the beginning of the course. Why is Newton's law a formula for acceleration, rather than something else. We will then be in a position to give a provisional answer to this question, or at least to relate it to another question, which in my view is one of the central mysteries of the universe.

The Cosmological Principle and the Hubble expansion

The "Cosmological Principle" is the assumption that when we average over sufficiently large scales, every place in the universe looks the same as every other place — there is no privileged position — no "center" of the universe; and that all directions are equivalent — the universe is isotropic. Ugh! This is Philosophy. And besides, I don't believe it. One of the nice things that has happened in recent years is that we can now at least imagine how the cosmological principle might arise approximately from more physical principles. We will come back to this later, and also discuss the experimental status of the cosmological principle. For now, as cosmologists have done for a long time, we will simply assume the cosmological principle to be true.

When astronomers look at distant galaxies, they see them, on the average, moving away from us — receding. It is believed that the recession velocity is roughly proportional to the distance from us, at least until we go out far enough that the recession velocities are not small compared to the speed of light. This proportionality of recession velocity to distance is called the Hubble law.

Let me give a very impressionist view of the observational evidence for this. The measurement of the recession velocity is not so difficult because astronomers can observe the Doppler shifts in spectral lines. Spectral lines are sharp peaks in the frequency spectrum of light (or other electromagnetic radiation). They are associated with quantum mechanical jumps between particular quantum states of atoms or molecules, and appear on top of the continuum spectrum emitted by a hot object. The frequencies of these lines are determined just by quantum mechanics, so we know what they are in the rest frame. Here is a spectrum for a quasar (which is presumably just a distant galaxy doing strange things before its center settled down) spectrum, from a preprint in 1996, W. Zheng, *et. al*, "A Composite HST Spectrum of Quasars," arXiv:astro-ph/9608198.

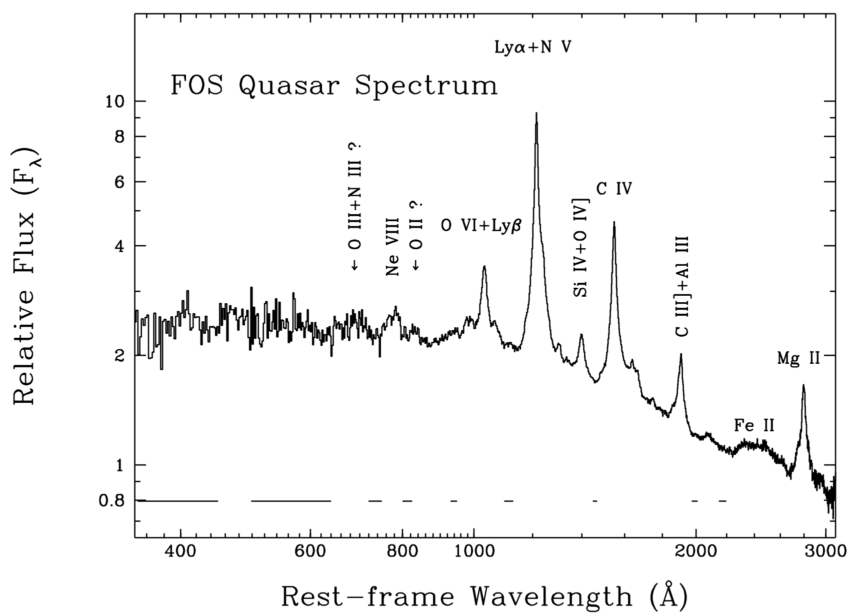
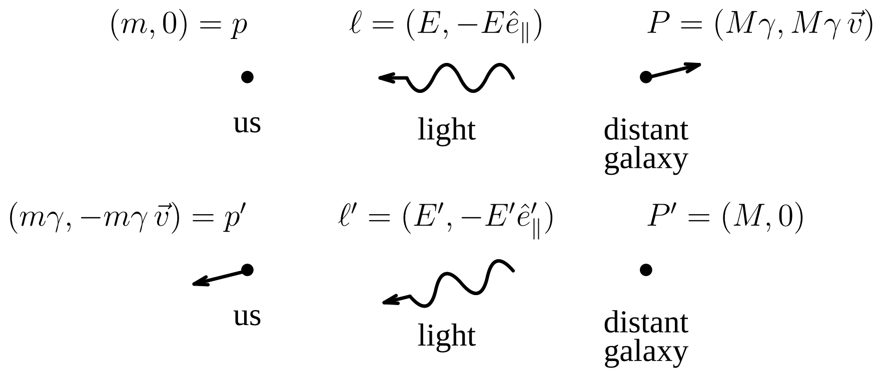


Figure 1: Composite FOS spectrum of 101 quasars, binned to 2°A. Prominent emission lines and the Lyman limit are labeled, and two possible emission features are marked. The continuum fitting windows are marked with the bars near the bottom.

To lowest order in the recession velocity in units with $c = 1$, the fractional Doppler red-shift in the light that reaches us from a receding object is just 1 minus the component of velocity directly away from us. Here is a quick derivation. The diagrams show light coming from a distant galaxy reaching us in two frames — our frame and the rest frame of the galaxy.



Also shown are the 4-momenta of an object on earth of mass m and an object in the galaxy of mass M.[1] Using the invariance of the invariant product

$$P \cdot \ell = M\gamma E - M\gamma\vec{v} \cdot (-E\hat{e}_\parallel) = ME\gamma(1 + v_\parallel) = ME' = P' \cdot \ell' \tag{1}$$

And therefore

$$E/E' = \frac{1}{\gamma(1 + v_\parallel)} \approx 1 - v_\parallel \tag{2}$$

This is just what we want to know to verify the Hubble law.

But the interpretation of the observations is not trivial because we do not have accurate ways of measuring the distances to far-away galaxies. For things that are nearby on an astronomical scale, we can measure the change in angle in the line of sight (compared to that for far away objects) as the earth moves around the sun, and use trigonometry to relate the distance of the object to the earth-sun distance (this is called parallax). Check out The ABC's of Distances for adiscussion of parallax in Ned Wright's Cosmology Tutorial.

But for more distant objects, we can only infer their distance if they are or they contain objects whose intrinsic brightness we think we know. Such things are called "standard candles." One of the keys to this business is to find good standard candles. If we see distant standard candles, and measure their observed brightness, we can calculate their distance by comparing the observed brightness with the intrinsic brightness. In principle this should work, but in practice, it is a much chancier proposition than using trigonometry. It will be accurate if we have a good theory of the intrinsic brightness and if the observed brightness is not affected by junk in the universe between the object and us. We will come back to this next week.

Hubble's original data (figure 2) was quite poor, but these days the Hubble law is known to work well out to distance of hundreds of megaparsecs, where a parsec is about 3.26 light years. Figure 3 shows some data using Type Ia Supernovas as a standard candle from Riess, Press and Kirshner (1996), quoted in Ned Wright's Cosmology Tutorial

Another reason that we believe the Hubble law is not experimental, but theoretical — it is the only possibility consistent with the cosmological principle. Let us see why this is so. We will do

[1]The masses will cancel out, and it would be more elegant to use the 4-velocity 4-vector, which is the 4-momentum divided by the mass — that since that is one more definition to remember, and we don't really need it, I have not spent much time discussing it.

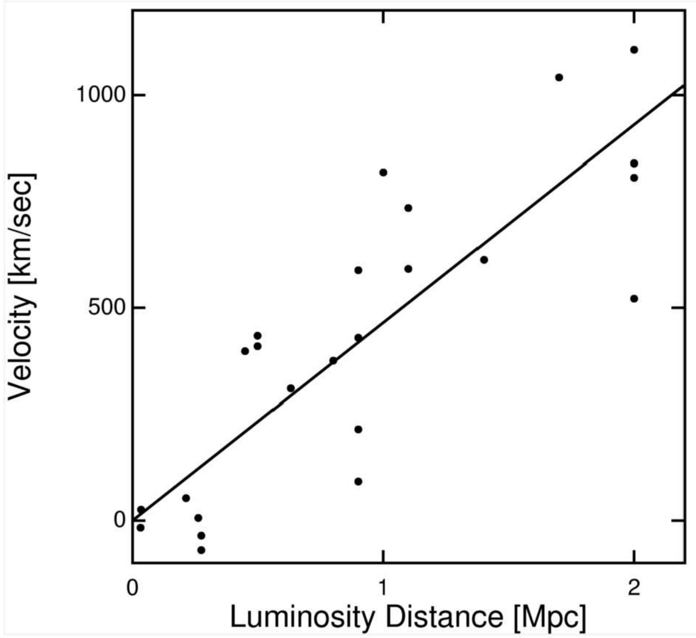


Figure 2: Original Hubble data (1996)

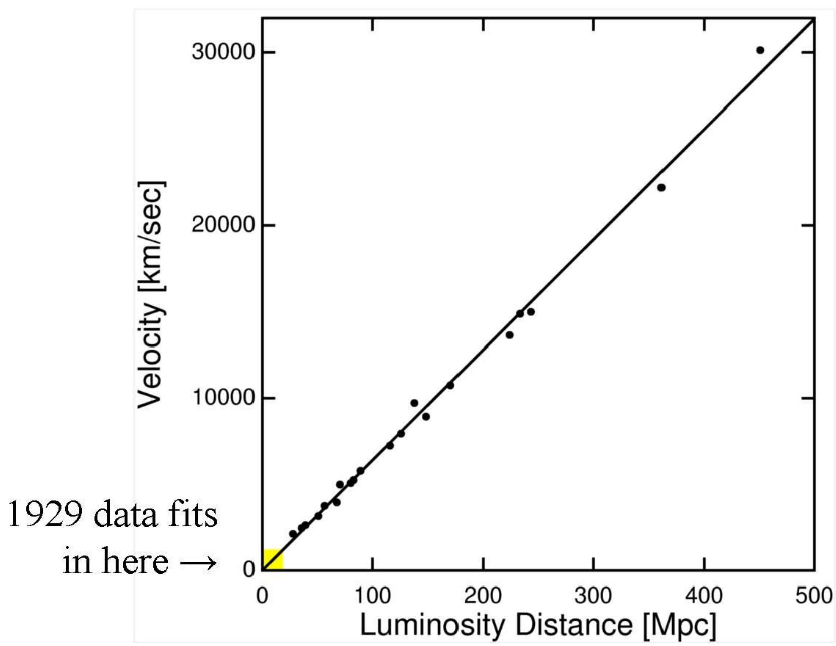


Figure 3: A Hubble plot using Type Ia SNe as a standard candle from Riess, Press and Kirshner (1996)

this assuming that the velocities are small compared to the speed of light — this will be valid in some relatively large region around the origin, as we will see. Then we can do the whole analysis at some fixed time t, and not worry about the changes in time that occur in Lorentz transformations.

The way the proof works is that we begin by writing the law that relates velocity and position in a particular frame, defined by having the galaxy at the origin of the coordinate system at rest. We look at a galaxy at position $\vec{r}_0$ and find that it is moving with velocity $\vec{v}_0$. Then we transform to a new frame in which this galaxy is at the origin and at rest, and we demand that the law has the same form.

Here is the proof. Suppose that at some time t there is some law that specifies the velocity of each galaxy in terms of its position. This law can be expressed in terms of a function $\vec{\mathcal{H}}(\vec{r})$ which gives the vector velocity $\vec{v}$ of the galaxy at position $\vec{r}$ in the inertial frame in which the galaxy at $\vec{r} = 0$ is at rest:

$$\vec{v} = \vec{\mathcal{H}}(\vec{r}) \tag{3}$$

Now suppose that we go an inertial frame in which the galaxy at $\vec{r}_0$ is at rest and go to a coordinate system in which this galaxy is at the origin. In the original coordinate system, this galaxy has velocity $\vec{v}_0 = \vec{\mathcal{H}}(\vec{r}_0)$. Thus in the new coordinate system, the positions and velocities are given by

$$\vec{r}\,' = \vec{r} - \vec{r}_0 \qquad\qquad \vec{v}\,' = \vec{v} - \vec{v}_0 = \vec{\mathcal{H}}(\vec{r}\,) - \vec{\mathcal{H}}(\vec{r}_0) \tag{4}$$

Now according to the cosmological principle, we must have the same law in the new coordinate system, so that

$$\vec{v}\,' = \vec{\mathcal{H}}(\vec{r}\,') \tag{5}$$

Putting (4) and (5) together, you see that $\vec{\mathcal{H}}$ must satisfy

$$\vec{\mathcal{H}}(\vec{r} - \vec{r}_0) = \vec{\mathcal{H}}(\vec{r}\,) - \vec{\mathcal{H}}(\vec{r}_0) \tag{6}$$

But this means that the function $\vec{\mathcal{H}}(\vec{r}\,)$ must be linear — each component of $\vec{\mathcal{H}}$ must be just a linear combination of the components of $\vec{r}$. This may be obvious, but let me belabor the point a little by giving you a "proof." Look at some component of $\vec{\mathcal{H}}$, say $\mathcal{H}_x$, and write (6) as

$$\mathcal{H}_x(\vec{r}_1 - \vec{r}_2) = \mathcal{H}_x(\vec{r}_1) - \mathcal{H}_x(\vec{r}_2) \tag{7}$$

Now differentiate both sides with respect to x_1 and x_2 (or y or z). On the right hand side, this gives zero because the two terms depend either on x_1 or x_2 but not both. The left hand side can be written entirely in terms of the derivative with respect to x_1 using the chain rule:

$$\frac{\partial^2}{\partial x_1 \partial x_2}\mathcal{H}_x(\vec{r}_1 - \vec{r}_2) = 0 \Rightarrow \tag{8}$$

$$\frac{\partial^2}{\partial x_1^2}\mathcal{H}_x(\vec{r}_1 - \vec{r}_2) = 0 \tag{9}$$

(9) is true for all $\vec{r}_2$, so we can set $\vec{r}_2 = 0$ and rename $\vec{r}_1 \to \vec{r}$, so

$$\frac{\partial^2}{\partial x^2}\mathcal{H}_x(\vec{r}) = 0 \tag{10}$$

This means that $\mathcal{H}_x(\vec{r})$ is at most linear in x. The same argument works for y and z. But setting $\vec{r}_1 = \vec{r}_2$ in (7) gives $\mathcal{H}_x(0) = 0$, and so we can write

$$\mathcal{H}_x(\vec{r}) = b_{xx}x + b_{xy}y + b_{xz}z \tag{11}$$

where the bs are constants. This works for the other components as well, so our proof is complete.

Now isotropy, the assumption that all directions are equivalent, implies that there is no special direction. This implies that $\vec{\mathcal{H}}(\vec{r})$ behaves like a vector under rotation. The only function of $\vec{r}$ that is linear, behaves like a vector, and does not involve any other fixed vector (which would be a special direction) is $\vec{r}$ itself, possibly multiplied by a constant. The constant may depend on the time t, but cannot depend on anything else. Thus

$$\text{The cosmological principle} \Rightarrow \vec{v} = H(t)\,\vec{r} \tag{12}$$

This is the Hubble Law! If we put ourselves at the origin (which doesn't matter because all points are equivalent) it says that a galaxy at $\vec{r}$ is moving with a velocity that is in the $\hat{r}$ direction — that is directly away from us — and with velocity proportional to $|\vec{r}|$. Notice that the function $H(t)$ has units of inverse time. The value of $H(t)$ today is called the Hubble constant, H_0. It is conventionally given in the rather ridiculous units of kilometers per second per megaparsec. It is sometimes further expressed in terms of a dimensionless quantity, h, defined by

$$H_0 \equiv 100\,h\,\frac{\text{km/s}}{\text{Mpc}} \tag{13}$$

where h is measured to be 0.71 ± 0.04. This means, for example, that for a galaxy at a distance of 100 Mpc, we expect a recession rate of 7.1×10^3 kilometers per second, that is about 2% of the speed of light. To convert these crazy units into an inverse time, we can convert all the distances to light years, and the time to years, using the fact that $c \approx 3 \times 10^5$ km/s=1 light year/year:

$$H_0 = \left(100\,h\,\frac{\text{km/s}}{\text{Mpc}}\right)\left(\frac{\text{1 light year/year}}{3 \times 10^5\,\text{km/s}}\right)\left(\frac{1}{10^6 \times 3.26\,\text{light years/Mpc}}\right) \approx \frac{h}{10^{10}}\,\frac{1}{\text{years}} \tag{14}$$

Which implies that $1/H_0$, which is called the Hubble time, is about 13.7 billion years. You can see on your PC what this looks like (though of course you can imagine) in the *Mathematica* file hubble.nb.

Equation (12) describes what is called an expansion of the universe (at least if $H(t)$ is positive, which it is today) because the distances between all pairs of galaxies grow in exactly the same way. It is as if the space between the galaxies is expanding at the rate $H(t)$. This is illustrated in the animation in hubble.nb. The Hubble time, $1/H_0$, is the time it would take for the universe to expand by a significant factor (e) if the time dependence of $H(t)$ is ignored. The Hubble time is thus a kind of approximate age of the universe — the time it has taken for the size of the universe to change a lot.[2] This gives rise to the notion of the observable universe, which is an imaginary

[2] To get a really accurate age, we have to understand how $H(t)$ depends on time, which we will discuss later.

sphere with us at the center and radius of about c/H_0 — that is 13.7 billion light years.[3] This radius is just the distance that light has traveled since the universe was much smaller than it is today. This also gives rise to the notion of the Big Bang. If we run the expansion backwards, we get to a point where things start to get complicated! We will talk more about this later.

Of course, the Hubble law, (12), and the cosmological principle from which it follows are is not really true. The universe is not really the same everywhere and the actual motions of galaxies are more complicated than (12). Galaxies come in clusters and superclusters in which they orbit around each others in complicated ways. But the idea is that this complicated local behavior is superimposed on the simple Hubble expansion. Below is a very very rough guide to the sizes of things:

The scales of the observable Universe	
~ 0.01 Mpc	visible size of large galaxies
~ 0.1 Mpc	dark size of large galaxies
~ 1 Mpc	distance between galaxies
~ 10 Mpc	big clusters bound by gravity
~ 100 Mpc	largest regions of enhanced density
$\updownarrow$ cosmological principle applies in this region???? $\updownarrow$	
$\sim 10^4$ Mpc	$\sim c/H_0$ - the observable universe

Newton's theorem and the expansion of the Universe

I'm sure you have all heard that general relativity implies that the universe might be curved, and might be finite, rather like the two dimensional space on the surface of a balloon. This would make it difficult for us to analyze it in purely Newtonian terms. Also, as we know only too well, Newton's picture of gravity cannot be exactly right because it does not account for special relativity. Nevertheless, at least if the universe is made of ordinary matter (an assumption that we will question in later lectures) we can use Newtonian gravity to learn something about the evolution of $H(t)$. We can do this because of the cosmological principle. Since the universe is (assumed to be) the same everywhere, we can analyze the evolution of $H(t)$ in a sufficiently small region that the velocities never get relativistic, and the curvature of space is not important. Then we can use Newtonian mechanics and Newtonian gravity to discuss not just the velocities of the galaxies, but

[3] mmm

the evolution of their velocities. Let's begin by adopting a coordinate system with us fixed at the center, and asking about the motion of a distant galaxy at $\vec{r}$.

This subject contains some subtleties. We will start with a deceptively simple analysis, and then go back and think about it. Consider the contribution to the forces on us and on the distant galaxy from a spherical shell of mass (galaxies) centered at the origin. Such a shell never makes any contribution to a force on us because of Newton's theorem. So there is no force on us and we remain fixed at the origin of the coordinate system. However, the distant galaxy feels a force from every shell with radius less than $r = |\vec{r}\,|$, pulling it towards us. The total force on the galaxy is towards us with magnitude

$$\frac{4\pi\, r^3}{3}\,\frac{G\, m\, \rho}{r^2} = \frac{4\pi}{3}\, G\, m\, \rho\, r \tag{15}$$

where m is the mass of the galaxy and ρ is the mass density of the universe at the time.

Now what does this mean? The obvious problem with the analysis is that it seems to depend on choosing the center of the spheres to be where we are. Why not choose the center some place else. If we do that, the force on us will not be zero, so how can we say that we are in an inertial frame? It is true that if we do this the galaxies are moving in some complicated way through our spheres, but that should make no difference to Newton's theorem because in Newtonian mechanics, the gravitational force depends only on the instantaneous positions of the various masses.

One problem here is one of infinity. If you think about the force on any galaxy in the universe from all of the others, in Newton's theory, the answer is simply undefined. Replacing the sum over galaxies by an integral over the position of the galaxies (which should be a good approximation for distances much larger than the typical distances between galaxies, or at least clusters) it is

$$\vec{F} = G\, m\, \rho \int d^3r'\, \frac{\vec{r}\,' - \vec{r}}{|\vec{r}\,' - \vec{r}\,|^3} \tag{16}$$

This integral is not well defined because it gets contributions from regions of the universe that are infinitely far away. It depends on exactly how we take the limit that extends the integral over all space. If we consider (16) as the limit of an integral over a large sphere as the radius of the sphere goes to infinity, then this reduces to the previous discussion, and as we have seen, the answer depends on where we choose the center of the spheres. In general, the limit depends on some complicated way on how one takes the volume to infinity. This is not useful. So what is actually going on?

First note that because of the equivalence principle, we can think about gravitational accelerations, rather than forces — this is nice because the mass of the galaxy cancels out.

$$\vec{a} = G\, \rho \int d^3r'\, \frac{\vec{r}\,' - \vec{r}}{|\vec{r}\,' - \vec{r}\,|^3} \tag{17}$$

However, this doesn't help with the infinite volume problem — we don't know the acceleration of a given galaxy any more than we know the force on a given galaxy.

But fortunately, we are not really interested in the acceleration of any particular galaxy. What we mean by a coordinate system with us fixed at the center is that the coordinate system is accelerated by the local gravitational field at our position. Thus the situation is the same here as with the tidal force. The actual acceleration that we care about on a distant galaxy is the difference between the gravitational acceleration on the galaxy and the gravitational acceleration on us.

$$\vec{a}_{\text{relative}} = G\,\rho \int d^3r' \, \frac{\vec{r}\,' - \vec{r}}{|\vec{r}\,' - \vec{r}\,|^3} - G\,\rho \int d^3r' \, \frac{\vec{r}\,' - \vec{r}_0}{|\vec{r}\,' - \vec{r}_0\,|^3} \tag{18}$$

In general relativity, we would simply say that we can go to a locally inertial frame in which there is no gravitational field where we are.

While the gravitational acceleration on a particular galaxy is not well defined, the **difference in the acceleration on any two galaxies** is much better defined. You still have to be careful about how you take the limit of infinite volume. But unless you do something dumb the difference makes sense. For example, it is easy to see that if we define the limit in (18) by taking a large sphere to infinity, the difference is independent of where we take the center of the sphere.

Since we computed (15) by adding up the contributions of spheres centered on our position, the acceleration of gravity on us computed in this way vanishes. Thus when we compute the acceleration on the distance galaxy in the same way, we are actually computing the difference between the acceleration of the galaxy and that of our own, which is what we want. Likewise, if we compute the difference, then it will not matter where we put the center of our spheres — we will always get (15). So, understood properly, the simple result is correct, though the argument is inadequate!

Hubble dynamics

We can now follow the distant galaxy as it moves with the expansion of the universe. The acceleration is

$$\ddot{\vec{r}} = -\vec{r}\,\frac{4\pi}{3}\,G\,\rho(t) \tag{19}$$

(19) is consistent with the Hubble law in the following sense. If at some time, the Hubble law, (3), is satisfied, the dynamics implied by (19) implies that it will be satisfied at subsequent times. To see this explicitly, note that (19) can be written as

$$\frac{d}{dt}\,\vec{v} = -\vec{r}\,\frac{4\pi}{3}\,G\,\rho(t) \tag{20}$$

Thus the rate of change of $\vec{v}$ is proportional to $\vec{r}$, so if $\vec{v}$ is proportional to $\vec{r}$ at some time, it stays that way.

To say more about the evolution of the expanding universe, it is convenient to rewrite (19) as follows:

$$\ddot{\vec{r}} = -\vec{r}\,\frac{4\pi}{3}\,G\,\rho = -\vec{r}\,\frac{G\,M}{r^3} \tag{21}$$

where

$$M = \frac{4\pi\,\rho\,r^3}{3} \tag{22}$$

is the mass inside an imaginary sphere of radius $r = |\vec{r}|$. That is M is the sum of the masses of all the stuff that is closer to us than the galaxy of interest. The advantage of writing things in terms of M is that M doesn't change as r changes. The size of the imaginary sphere changes as the galaxy we are looking at moves farther way, but on the average, because all parts of the universe are expanding at the same rate, stuff doesn't cross the surface of the imaginary sphere. So M remains constant. This is illustrated in the *Mathematica* file hubble.nb it you set $sphere = 1$.

Taking the dot product of both sides of (21) by $\dot{\vec{r}}$, we can write the left hand side as

$$\dot{\vec{r}} \cdot \ddot{\vec{r}} = \frac{d}{dt} \frac{1}{2} \left(\dot{\vec{r}}\right)^2 \tag{23}$$

and the right hand side as

$$-\dot{\vec{r}} \cdot \vec{r} \frac{G\,M}{r^3} = -\dot{\vec{r}} \cdot \frac{\vec{r}}{r} \frac{G\,M}{r^2} = \dot{\vec{r}} \cdot \frac{\partial r}{\partial \vec{r}} \frac{\partial}{\partial r} \frac{G\,M}{r} = \dot{\vec{r}} \cdot \frac{\partial}{\partial \vec{r}} \frac{G\,M}{r} = \frac{d}{dt} \frac{G\,M}{r} \tag{24}$$

or

$$\frac{d}{dt} \frac{1}{2} \left(\dot{\vec{r}}\right)^2 = \frac{d}{dt} \frac{G\,M}{r} \tag{25}$$

which implies that

$$\frac{1}{2} \left(\dot{\vec{r}}\right)^2 - \frac{G\,M}{r} = \text{constant} \equiv \frac{1}{2} C \tag{26}$$

or

$$\frac{1}{2} (\vec{v}\,)^2 - \frac{4\pi\,G\,\rho}{3} (\vec{r}\,)^2 = \frac{1}{2} C \tag{27}$$

The constant C in (26) and (27) depends on the initial conditions. There is a theoretical prejudice today in favor of $C = 0$, for reasons that I will try to explain later.[4] This means that the universe lives in a flat space, which makes it somewhat easier to understand. However, it is still possible that $C \neq 0$, in which case general relativity implies that the universe is curved. We won't talk about this in detail, except to note that if C is negative, the universe is finite and positively curved, like the three dimensional analog of the surface of a sphere, while if C is positive, the universe is negatively curved and infinite. We will assume that $C = 0$, and just briefly discuss what happens in the other cases. If we ignore relativity, the physical significance of C can be easily seen from (26). If $C > 0$, the expansion of the universe will go on forever, and as the distance to the distant galaxy r goes to ∞, the speed of the galaxy will go to a finite nonzero limit. If $C < 0$, there will be a largest value of r at which the galaxy stops, and gravitational attraction turns the expansion of the universe into contraction. If $C = 0$, the balance of kinetic energy and gravitational attraction is such that the speed of the distant galaxy goes to zero as $r \to \infty$, but it never stops and turns around.

Dynamics of a flat universe

Using the Hubble law, (3), in (27) gives

$$\frac{1}{2} H(t)^2 (\vec{r}\,)^2 - \frac{4\pi\,G\,\rho}{3} (\vec{r}\,)^2 = \frac{1}{2} C \tag{28}$$

[4]In any event, there is a sense in which the universe is remarkably close to $C = 0$.

Taking $C = 0$ in (28), dividing both sides by r^2 and putting in the time dependence of $\rho(t)$ explicitly, we find

$$H(t)^2 = \frac{8\pi\, G\, \rho(t)}{3} \tag{29}$$

The density implied by (29),

$$\rho_c(t) = \frac{3H(t)^2}{8\pi G} \tag{30}$$

is called the critical density. Because both $\rho(t)$ and $H(t)$ are measurable, at least in principle, (29) is a test of the flatness of the universe (assuming that the rest of our assumptions are correct). If the measured density is greater than critical, $\rho(t) > \rho_c(t)$, then C in (27) is negative, and the universe is closed and positively curved. If the measured density is less than critical, $\rho(t) < \rho_c(t)$, then C in (27) is positive, and the universe is infinite and negatively curved.

The matter that we can actually see in the universe contributes only a very small fraction of the critical density. However, this doesn't mean that the universe is not flat.

Newton's theorem and dark matter

There is a lot of evidence that galaxies contain much more matter than we can see. What we see we see because it emits light — it is called luminous matter — mostly stars or gas clouds. The most convincing evidence that there is a lot of other stuff comes from studies of the orbital velocity of gas far from the galactic centers, beyond almost all of the visible matter. What one sees is that the speed at which things orbits around the center of the galaxy **remains roughly constant with distance far beyond the radius at which you see significant visible matter.** Newton says that for a circular orbit at radius r in a spherically symmetric collection of mass,

$$\frac{v^2}{r} = \frac{G\, M(r)}{r^2} \tag{31}$$

where $M(r)$ is the mass inside a sphere of radius r. Thus one simple way of explaining the constancy of the orbital velocity is to assume that there is a large spherically symmetric halo of dark matter around each galaxy with a density that goes approximately like $1/r^2$, where r is the radius. Then the amount of matter contained in a sphere of radius r is

$$M(r) = \int_0^r 4\pi {r'}^2\, dr'\, \rho(r') \propto \int_0^r dr' \propto r \tag{32}$$

which cancels a $1/r$ on the right hand side of (31), so that the rotation speed v remains constant. This much of the subject you should understand.

Whatever it is that contributes this mass density proportional to $1/r^2$ (except for the stars and gas that we see) is called dark matter. So what is this stuff? Is it real? Is it matter that we know about or something else?

Recently, another interesting way of looking at the effect of dark matter has emerged. The rotation speed arguments we have just talked about involve the effect of the dark matter's gravity on ordinary matter. This has more-or-less convincingly established that there is dark matter associated with the luminous matter in galaxies. There are also indications that there is more dark matter associated with clusters of galaxies. Weird as it sounds, one can also see the effect of dark matter's gravity on light. The idea here is to use so-called "gravitational lensing." Because light can be bent by a gravitational field, it is possible that a massive cluster of galaxies could bend the light from an even more distant galaxy (or quasar) so that we on earth would see the light coming at us from several directions.[5] This has actually been seen, first in radio-telescope observations, and more recently and spectacularly in Hubble Space Telescope images and other visual data. Figure 4 shows one such image. In this Hubble photo, around the bright oval galaxies of a distant cluster,

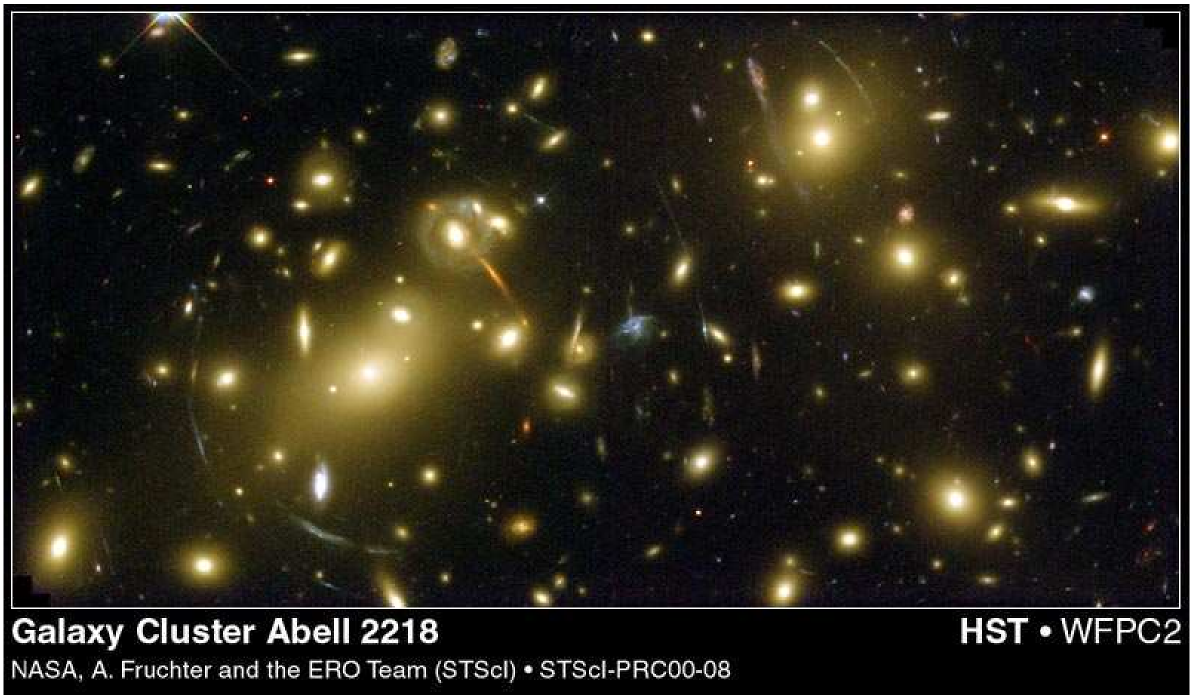


Figure 4: Hubble images showing gravitational lensing.

we see the wispy arcs of a still more distant object, focused and distorted by the gravitational lens produced by the dark matter of the cluster. One can use this phenomenon to study the distribution of the dark matter. Figure 5 dramatically shows the result of such a study of another cluster of galaxies - a plot of the matter density in a cluster, showing high peaks at the visible galaxies, but an enormous background of dark matter between them.[6]

If the dark matter really is some kind of matter, there are indirect but fairly convincing arguments that it cannot be made out of normal stuff. The evidence comes from the fact that we have a picture of how light nuclei are formed in the very early universe that works reasonably well, and a lot more normal matter would mess up this nice picture (of what is called "primordial nucleosyn-

[5] or even, if the geometry is just right, from an "Einstein ring."

[6] from a Bell-Labs new release in 1997.

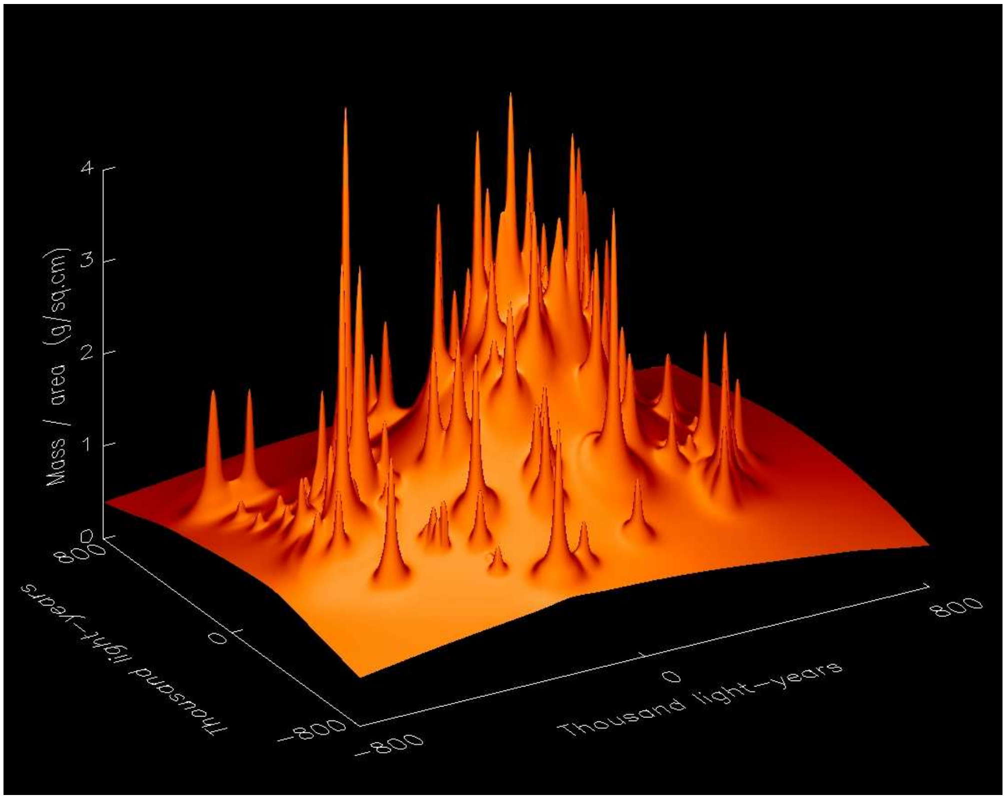


Figure 5: Mass distribution in a cluster of galaxies showing dark matter bakground.

thesis" — for example, if there were enough baryons to account for all the dark matter, the theory predicts that there would be much less deuterium in the universe than we actually see!). So most people assume that it is some kind of new particle! There are many ongoing experimental efforts to see dark matter particles as the earth moves through the dark-matter halo of our galaxie. At least one experimental group believes that that they have seen something, but the data so far is not convincing, nor does it tell us much about what this stuff might be.

I am not going to talk about all this in detail. Here is a link to reviews on the web:

astron.berkeley.edu/∼mwhite/darkmatter/dm.html — A brief but up-to-date introduction with additional links.

lecture 24

Topics:
- Where are we?
- Running Hubble backwards
- Relativistic cosmology
- Back to the big bang
- Thermal equilibrium
- The hot bang and the CMBR
- Temperature and phases

Where are we?

After our introduction to the Hubble expansion in the last lecture, we are now going to work our way back towards the beginning of the universe. This is sort of fun, for its own sake, but it will also, I claim, get us closer to an answer to the question we asked at the very beginning of the course — why is Newton's third law a formula for acceleration.

Running Hubble backwards

Let us start by going back to the laws we derived to describe the evolution of the Hubble expansion. If we put ourselves at the origin and take $\vec{r}$ to be the position of some distant galaxy, the Hubble law can be written as

$$\dot{\vec{r}} = H(t)\,\vec{r} \tag{1}$$

where $H(t)$ is the Hubble parameter. Then, assuming that we can use Newton's theory of gravity and ignore relativity, we found that gravity slows the Hubble expansion

$$\ddot{\vec{r}} = -\vec{r}\,\frac{4\pi}{3}\,G\,\rho(t) \tag{2}$$

Then by considering the conservation of the mass inside a sphere with radius $|\vec{r}\,|$ (still ignoring relativity), we found that we could integrate (2) to obtain

$$\frac{1}{2}(\vec{v}\,)^2 - \frac{4\pi\,G\,\rho(\vec{r}\,)^2}{3} = \frac{1}{2}C \tag{3}$$

where the constant C depends on the initial conditions. Again, this derivation was valid for a universe dominated by slowly moving matter, so long as the cosmological principle is satisfied.

It is conventional and convenient to get rid of the vectors in (1)-(3). We can do this by defining an (arbitrary) distance scale a between points in the universe. For example, we could take $a = |\vec{r}\,|$, the distance between us and the distant galaxy. But the notion is more general, as it has to be because we want to use it to describe the universe at earlier times, before we or even galaxies were around. a is just an arbitrary scale that measures the relative size of the universe. The value of a doesn't matter at all, but if a doubles, that means that the distance between (far apart) things in the universe has doubled!

Using $a = |\vec{r}\,|$ and that fact that the direction of $\vec{r}$ doesn't change with time, we can write

$$\vec{r} = a\,\hat{r} \qquad \dot{\vec{r}} = \dot{a}\,\hat{r} \qquad \ddot{\vec{r}} = \ddot{a}\,\hat{r} \tag{4}$$

and rewrite (1)-(3) as

$$\frac{\dot{a}}{a} = H(t) \tag{5}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3} G\,\rho(t) \tag{6}$$

and

$$\frac{\dot{a}^2}{a^2} = H^2 = \frac{8\pi\,G\,\rho}{3} + \frac{C}{a^2} \tag{7}$$

Now for simplicity, let us assume that $C = 0$. As I have said before, there are theoretical reasons (that we will discuss next time) to think that this is a good approximation, and some observational support for the detailed picture that emerges. We can then follow the Hubble evolution backwards towards the big bang, at least for a while.

For $C = 0$, we can take the square root of (7) and write

$$\frac{\dot{a}}{a} = \sqrt{\frac{8\pi\,G\,\rho}{3}} \tag{8}$$

Note that have taken the positive sign, corresponding to expansion.

To determine the time evolution of a, we need a relation between a and ρ. This relation depends on what kind of stuff our universe is made of at the time of interest. For slowly moving matter, we can simply use the fact $\rho\,a^3$ is a constant, because the density is inversely proportional to the volume. Thus

$$0 = \frac{d}{dt}(\rho\,a^3) = \dot{\rho}\,a^3 + 3\rho\,a^2\,\dot{a} \quad \Rightarrow \quad \frac{\dot{a}}{a} = -\frac{1}{3}\frac{\dot{\rho}}{\rho} \quad \text{or} \quad \dot{\rho} = -3\rho\dot{a}/a \tag{9}$$

Thus we can write (8) as

$$-\frac{1}{3}\frac{\dot{\rho}}{\rho} = \sqrt{\frac{8\pi\,G\,\rho}{3}} \tag{10}$$

This is a differential equation that we can solve for ρ (or course, the result is only meaningful as long as our assumptions are valid):

$$\dot{\rho} = -\sqrt{24\pi\,G}\,\rho^{3/2} \tag{11}$$

$$\frac{\rho^{-1/2}}{-1/2} = \int \frac{d\rho}{\rho^{3/2}} = -\sqrt{24\pi\,G}\int dt = -\sqrt{24\pi\,G}\,(t - t_0) \tag{12}$$

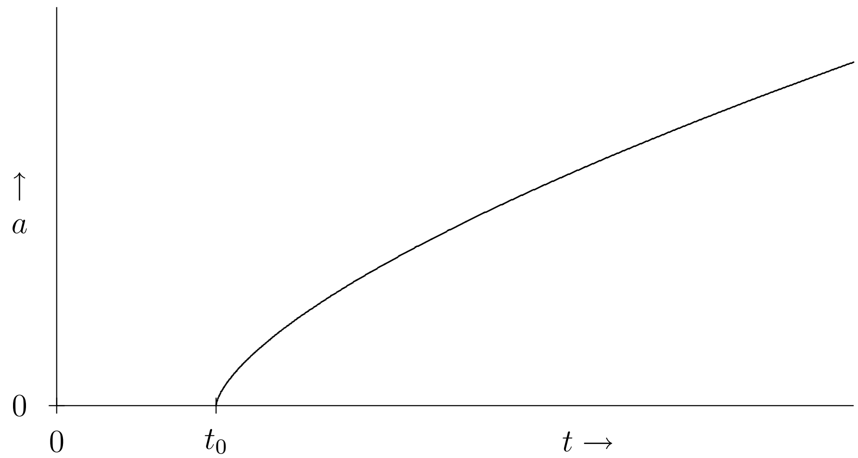
$$\rho^{-1/2} = \sqrt{6\pi\,G}\,(t - t_0) \tag{13}$$

$$\rho = \frac{1}{6\pi\,G\,(t - t_0)^2} \tag{14}$$

where t_0 is a constant set by initial conditions. As expected, the density decreases as the universe expands and increases as we run the tape backwards. Since $\rho\,a^3$ remains constant as t increases, (14) implies that a grows as

$$a \propto (t - t_0)^{2/3} \tag{15}$$

So if this were the whole story, the expansion of the universe would look something like this



To recapitulate, two separate pieces of physics are required to understand how the scale a evolves with time. One is the is the relation (7) (or (8) for $C = 0$) that describes the effect of gravity on the expansion. The other is the relation (9) that describes how the density ρ changes with changes in the scale factor a. A third relation, (6), is not independent of these two. We can derive it from (7) and (9). One way to do this is to multiply (7) by a^2 and differentiate (going backwards through the steps we used to get (7)) to get

$$2\dot{a}\ddot{a} = \frac{8\pi G}{3}(\dot{\rho}a^2 + 2\rho a\dot{a}) = \frac{8\pi G}{3}(-3\rho a\dot{a} + 2\rho a\dot{a}) = -\frac{8\pi G\rho a\dot{a}}{3} \tag{16}$$

Dividing by $2a\dot{a}$ then gives (6).

Relativistic cosmology

The simple behavior implied by (14) and (15) is not consistent with the universe we actually observe. A couple of important things have been left out. One of these has been known for a long time. There is good evidence that the big bang was hot as well as dense. Heat complicates matters in a couple of ways. Heat is the random motion of the constituents of matter. The higher the temperature, the faster the motion. But once stuff is moving around, we have not only density, but also pressure — we left pressure out of our simple analysis of the Hubble dynamics. Furthermore, as we go back farther in time, the temperature gets higher and higher, and eventually, the random motion of the particles approaches the speed of light. Then we cannot ignore relativity! In this regime, as we know from our study of relativity, all sorts of bizarre things happen. Particles are created and destroyed. Energy and momentum are conserved, but not mass. We must be careful.

A complete understanding of what happens when things get relativistic requires that we generalize Newton's theory of gravity to incorporate special relativity. The resulting theory is Einstein's general relativity. We are not going to discuss this here, except for one result, which will be all we need. It turns out that (7) remains correct in general relativity if ρ is interpreted not as the mass density, but as the relativistic energy density divided by c^2, or just the energy density in sensible units in which $c = 1$. This makes sense in that if we go back to low temperatures and things come to rest, the energy density just becomes the mass density. But in the relativistic regime, it just turns out that gravity affects all forms of energy, not just mass (that is rest energy), so we have to use the energy density instead in (7). It is worth writing (7) again, now that we know how general it is:

$$\frac{\dot{a}^2}{a^2} = H^2 = \frac{8\pi\, G\, \rho}{3} + \frac{C}{a^2} \tag{17}$$

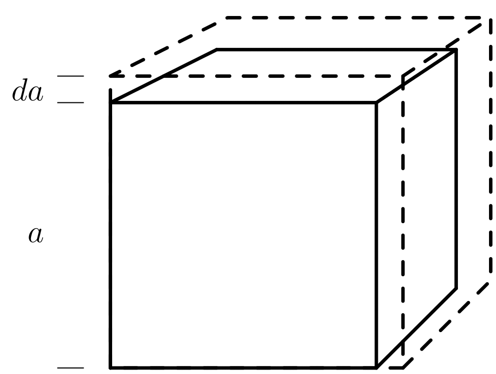
This is called the Friedmann equation. The parameter C here is related to the curvature of the space of the universe. $C = 0$ corresponds to flat space.

Let's see what becomes of (9) and (6) in the presence of pressure and relativity. First consider (9). To think about how this changes, let's consider the total energy in a cube with side a, and see how it changes with time in an expanding universe. Remember that a here is some distance that changes along with the space of the universe. This cube is like the sphere in **hubble.nb** except that now everything is bouncing around randomly and bumping into everything. On the average, no stuff comes into or goes out of the cube, or at least as much stuff comes in as goes out. We could think of it as having impermeable sides or pistons on all sides or something like that, so that collisions with the sides take the place of particles going in and out of the cube. This isn't quite right for any particular collision, but it should work for averages. And since temperature and pressure are the result of random processes, averaging over a very large number of individual processes, this is really what we want. Because we are now intepreting ρ as the energy density, the energy is $\rho\, a^3$. Now in this picture with the impermeable sides, because of the pressure, the energy changes with the volume — the pressure p does work on the boundary, which reduces the energy.

Let's analyze this quantitatively in two ways - first for the cube, and then for a region of arbitrary shape. Both will give the same result.

It is easiest to see what happens for the cube if we put one corner at the origin, so this corner doesn't move with the expansion. Then, as shown in the figure below, when $a \to a + da$ the three sides that do not touch the origin each move out a distance da. The area of each face is a^2 (it doesn't matter whether we use a^2 or $(a + da)^2$ or something in between because the difference is infinitesimal). Thus the force on the face is $p\, a^2$ and the work done on each face is $p\, a^2\, da$. Because there are three faces, the energy lost is $3p\, a^2\, da$, which is the change in energy in the cube. Then because $3a^2\, da = d(a^3)$, we can write

$$d(\rho\, a^3) = -3p\, a^2\, da = -p\, d(a^3) \tag{18}$$



Now let us do the analysis for a region of arbitrary shape. The work done by the pressure in an infinitesimal change of the boundary is a surface integral over the boundary of the force $d\vec{F}$ dotted into the change in the position of the boundary, $d\vec{b}$. The force of pressure has magnitude equal to the area dA times pressure p, and its direction is normal to the boundary. Thus the work done is the pressure times the area times the perpendicular motion of boundary, which is equal to p times the change in the volume —

$$dE = \int_A d\vec{F} \cdot d\vec{b} = \int_A p\, dA\, db_\perp = p\, dV \tag{19}$$

Thus here because $V = a^3$, (19) implies

$$\frac{d}{dt}(\rho\, a^3) = -p\, \frac{d}{dt}(a^3) \tag{20}$$

which agrees with (18).

(20) implies

$$\dot\rho\, a^3 + 3\rho\, \dot a\, a^2 = -3p\, \dot a\, a^2 \tag{21}$$

or

$$\dot\rho = -3\frac{\dot a}{a}(\rho + p) \tag{22}$$

Notice that this reduces to (9) when the pressure vanishes, as it should.

Using (22), we can find the appropriate generalization of (6) by following the steps we used in (16). Multiply (17) by a^2, and differentiate with respect to time.

$$\frac{d}{dt}\dot a^2 = \frac{d}{dt}\frac{8\pi\, G\, \rho\, a^2}{3} \tag{23}$$

$$2\dot a \ddot a = \frac{8\pi\, G}{3}\left(\dot\rho\, a^2 + 2\rho\, \dot a\, a\right) = \frac{8\pi\, G}{3}\left(-3(\rho + p)\, \dot a\, a + 2\rho\, \dot a\, a\right) \tag{24}$$

Dividing both sides by $2a\dot a$ gives

$$\frac{\ddot a}{a} = -\frac{4\pi\, G}{3}(\rho + 3p) \tag{25}$$

This is the relativistic generalization of (6), to which it reduces when $p = 0$ and ρ is the rest energy. Note that it is not at all obvious why it is (7) that does not change when we go to the relativistic limit, while (6) gets generalized to (25). This requires knowing that (7), with ρ reinterpreted at the energy density, is the correct general relativistic result.

Back to the big bang

Armed with the results of the previous section, we can go back a bit further into the history of the universe. But now that we have included the effect of pressure, we also have to understand the relation between the energy density ρ and the pressure p for relativistic stuff. An important example of relativistic stuff is a gas of photons — particles of light — radiation. We can compute the pressure in any shape container, so consider the pressure of a gas of photons with energy density ρ in a cubical container of side a. The pressure of such a gas arises because the photons bounce off the sides of the container. Suppose the sides are lined up with the coordinate axes. A photon with energy E and momentum $\vec p$, if all the components of $\vec p$ are nonzero, bounces around off all the sides. Each time it hits a side, the component of momentum perpendicular to the side changes sign, but the others remain unchanged. Consider the force on the sides perpendicular to $\hat x$. When the photon bounces off this side, it imparts an impulse $2|p_x|$. The time it takes to get back to the same side, because it must go a distance $2a$ in the x direction, is $2a/|v_x|$, where v_x is the x component of the velocity, which is p_x/E. The contribution of this photon to the impulse per unit time, which is the force, is thus

$$\frac{2|p_x|}{2a/|v_x|} = \frac{E}{a}\, v_x^2 \tag{26}$$

The contribution to the pressure is the force (26) divided by the area a^2. Then we get the total pressure by summing over all the photons

$$p = \sum \frac{E}{a^3}\, v_x^2 \tag{27}$$

But the pressure is the same on each side, so because the velocity of each photon is 1, the result must be

$$p = \frac{1}{3} \sum \frac{E}{a^3}\, (v_x^2 + v_y^2 + v_z^2) = \frac{1}{3} \sum \frac{E}{a^3} = \frac{1}{3}\, \rho \tag{28}$$

The pressure of a relativistic gas is 1/3 the energy density (in relativistic units, with $c = 1$, of course).

Now let us apply (28) to understand how the energy density of a gas of photons changes with the Hubble expansion. For $p = \rho/3$, (22) becomes

$$\dot\rho = -4\, \frac{\dot a}{a}\, \rho \tag{29}$$

This implies that $\rho\, a^4$ is constant. Thus the energy density of a relativistic gas falls like a^{-4} as the universe expands, faster than the energy density of nonrelativistic matter, which falls like a^{-3}. It may help to understand this result to realize that this means that the energies of the individual relativistic particles must be falling like $1/a$, because their number density clearly falls like a^{-3}. In the expanding box derivation above, the energies fall because they lose energy in the their collisions with the sides of the box. In the universe, there isn't any box, but if the relativistic particles are in thermal equilibrium, they lose energy bouncing off the other particles around them which are expanding away from them (just like the sides of the box).

An interesting thing about this is that the relation (29) remains true even when the density of the relativistic particles becomes so low that they no longer bounce off one another very often. But even in that case, (29) is satisfied in an expanding universe. What is happening is the that energies are effectively red-shifted down as the universe expands. The larger the universe is, the farther away the relativistic particles reaching some particular point are coming from. But the farther away they came from, the more red-shifted they are (because of the Hubble expansion). If the particles have mass, the process eventually stops when the particles become nonrelativistic. But for photons, and other massless particles, it continues forever and (29) remains true as the density goes to zero.

Furthermore, going in the other direction, this means that if we follow the history of a relativistic gas back in time, it not only gets more dense as we go back, but because the energies of the individual particles are increasing, the temperature increases as well. It turns out (as we will see below) that the average energy is proportional to the temperature, so the temperature of the relativistic gas goes like $1/a$.

Thermal equilibrium

There is one more component to the now standard picture of the hot big bang. One assumes that at some time in the early history of the universe, all of the particles were in thermal equilibrium at an enormously high temperature. What thermal equilibrium means is that the particles collide frequently enough that their motions are thoroughly randomized. You might think that this randomness would make it hard to understand how such a hot universe works. But in fact, exactly the opposite is true. In thermal equilibrium, all the important properties are determined on the average by a single parameter — the temperature. You can them follow the subsequent evolution of the universe, at least on the average, using the tools we have developed above. This seemingly paradoxical situation is beautifully explained in one of the best popular science books I know of, Steven Weinberg's **The First Three Minutes**, which I recommend for any of you who have not already read it. You will also learn (much) more about the connection between temperature and randomness if you take Physics 181. But it is so beautiful that I want to take time out from our study of the universe and talk just a little bit about thermal equilibrium and randomness.

A beautiful area of physics called "statistical mechanics" is the study of how random motion of particles is related to classical notions like heat and temperature. "Thermal equilibrium" is one of the basic concepts here. The idea is that randomness allows us to use probalistic arguments to understand the physics. What makes this powerful is that the number of particles is very large.

I will give you one example of the power of this approach — the Boltzmann distribution. Suppose that a system with a fixed, very large number N of degrees of freedom is in thermal equilibrium. The degrees of freedom may be the components of the position of single particles, or they may include additional coordinates like the orientation of molecules. It doesn't matter what they are. But we will assume that these degrees of freedom are essentially free between collisions, but that collisions happen frequently and randomly, so that each degree of freedom that is in thermal equilibrium is thoroughly randomized with all the others.

The coordinate for each degree of freedom $\mathcal{Q}_j$ has a corresponding generalized momentum $\mathcal{P}_j$. The only important facts that goes into the Boltzmann distribututions are the thermal randomness that makes each degree of freedom the same on the average, the fact that N is very large, and the fact that there is a conserved energy that is quadratic in each of the momenta. It is easiest to see what is going on if we choose the normalizations of the momenta so that the energy is just the sum of the squares,

$$E_{\text{tot}} = \sum_{j=1}^{N} \mathcal{P}_j^2 \tag{30}$$

For example, for an ordinary space momentum, this just means absorbing a factor of $1/\sqrt{2m}$ so that

$$\mathcal{P}_j = \frac{1}{\sqrt{2m}}\, p_j \tag{31}$$

This just makes the derivation easier because we don't have to carry around the extra factors.

Now the statement of the Boltzmann distribution is that the probability of finding a degree of freedom with momentum between $\mathcal{P}$ and $\mathcal{P} + d\mathcal{P}$ is completely calculable in terms of the temperature —

$$B(\mathcal{P})\, d\mathcal{P} = \frac{d\mathcal{P}}{\sqrt{2\pi\, E_{\text{av}}}}\, e^{-\frac{\mathcal{P}^2}{2E_{\text{av}}}} = \frac{d\mathcal{P}}{\sqrt{\pi kT}}\, e^{-\frac{\mathcal{P}^2}{kT}} \tag{32}$$

The function

$$B(\mathcal{P}) = \frac{1}{\sqrt{2\pi\, E_{\text{av}}}}\, e^{-\frac{\mathcal{P}^2}{2E_{\text{av}}}} = \frac{1}{\sqrt{\pi kT}}\, e^{-\frac{\mathcal{P}^2}{kT}} \tag{33}$$

is the Boltzmann distribution, where E_{av} is the average energy per degree of freedom which is related to the temperature by Boltzmann's constant,

$$k = \text{Boltzmann constant} = 1.38 \times 10^{-23}\ \text{J/°K} = 8.62 \times 10^{-5}\ \text{eV/°K} \tag{34}$$

according to

$$E_{\text{av}} = \frac{kT}{2} \tag{35}$$

It is easy to understand the exponential factor in (33). The key to the probability argument is simply conservation of energy, (30). Because of (30), any particular degree of freedom can have a generalized momentum anywhere between $-\mathcal{P}_{\text{max}}$ and $\mathcal{P}_{\text{max}}$ where

$$\mathcal{P}_{\text{max}} = \sqrt{E_{\text{tot}}} \tag{36}$$

But if one degree of freedom has generalized momentum $\mathcal{P}_1$, then the maximum possible value of **all** the others is reduced to

$$\mathcal{P}'_{\text{max}} = \sqrt{E_{\text{tot}} - \mathcal{P}_1^2} = \sqrt{\mathcal{P}_{\text{max}}^2 - \mathcal{P}_1^2} \tag{37}$$

This is the basic physics. It is unlikely to find one degree of freedom with a very large generalized momentum $\mathcal{P}_1$ because conservation of energy would then restrict the possible momentum values for the very large number of other degrees of freedom. Because the possible range of each of the other degrees of freedom is reduced by a factor of

$$\frac{\mathcal{P}'_{\text{max}}}{\mathcal{P}_{\text{max}}} = \sqrt{1 - \mathcal{P}_1^2/\mathcal{P}_{\text{max}}^2} \tag{38}$$

the probability of finding a value $\mathcal{P}_1$ for one degree of freedom should contain a factor of

$$\left(\frac{\mathcal{P}'_{\text{max}}}{\mathcal{P}_{\text{max}}}\right)^{N-1} = \left(1 - \mathcal{P}_1^2/\mathcal{P}_{\text{max}}^2\right)^{(N-1)/2} \approx \left(1 - \mathcal{P}_1^2/\mathcal{P}_{\text{max}}^2\right)^{N/2} \tag{39}$$

The limit that we are interested is N very large with the total energy growing linearly with N so that the average energy per particle is fixed.

$$\mathcal{P}_{\text{max}}^2 = E_{\text{tot}} = N\, E_{\text{av}} \tag{40}$$

so (39) becomes

$$\left(1 - \frac{\mathcal{P}_1^2}{N\, E_{\text{av}}}\right)^{N/2} \tag{41}$$

In the limit of very large N, the exponential factor (41) overwhelms any other dependence on the particular degree of freedom and gives the exponential in the Boltzmann distribution:

$$\lim_{N\to\infty} \left(1 - \frac{\mathcal{P}_1^2}{N\, E_{\text{av}}}\right)^{N/2} = e^{-\frac{\mathcal{P}_1^2}{2E_{\text{av}}}} \equiv e^{-\frac{\mathcal{P}_1^2}{kT}} \tag{42}$$

The rest is just a normalization factor to make the total probability equal to one. You can see from this argument that the factor of $1/2$ in the relation between $E_{\rm av}$ and kT, (35), comes from the fact that the energy is a quadratic function of $\mathcal{P}$.

This is illustrated in the *Mathematica* worksheet boltzmann-dynamic.nb. Here we start with a uniform random distribution of the momenta between -1 and 1 (in some arbitrary units) of 5000 particles moving in 2 dimensions. The momenta are then randomized by a series of collisions, each conserving energy and momentum. Particle 1 collides with 2, then 2 with 3 and so on to 1000. The whole process is then repeated 10 times and you can watch as the momenta get randomized. You can see dramatically how the initial uniform distribution gets converted in a Boltzmann distribution by the power of probability.

The hot bang and the CMBR

Let's assume that the hot big bang picture is correct, and think about what the universe looked like when it was very hot. First of all, normal matter, made of neutral atoms certainly didn't exist, because the high energy collisions would have completely ionized all the atoms. The universe would be a plasma. Furthermore, since particle number is not conserved in relativistic collisions, particles and their antiparticles can be produced and destroyed. So for example, while it seems that in the universe today, there are a lot more electrons than there were positrons, long ago when the universe was less than a millionth its current size, there were almost as many positrons as electrons. The small excess of electrons that eventually became the electrons in our atoms was quite unimportant at early times. For the same reason, at even earlier times, there was a lot of other stuff around in the early moments of the universe that we don't see much of today — heavy unstable particles which today we can make only at large accelerator laboratories and which quickly decay back into ordinary stuff were as common in the very early universe as electrons. These heavy particles disappear when the universe cools to a temperature such that the typical particle energy is below their mass. It is all quite strange — but simple, in a funny way, because everything is more or less fixed just by the temperature.

Now why would anyone believe this? We cannot, after all, go back and do experiments on the early universe. Why is this discussion science? The answer is that we can almost see it! At least we can look back toward the beginning of the universe by looking far away in the universe, because the light from far away regions of the universe has taken a long time to get to us. But we can't look back all the way. Once the universe gets so hot that atoms dissociate into ions, photons cannot go very far without colliding with electrons — the universe becomes opaque. Thinking about this in the other direction is even more interesting. As the universe cools to below the temperature at which atoms dissociate (a few thousand degrees C), it becomes transparent to photons, which means that photons fall out of thermal equilibrium. From then on, most of the photons just move freely, never colliding with anything again. This "gas" of photons continues to behave like relativistic stuff, while the atoms are nonrelativistic. Thus as the universe continues to expand, the energy density in the photons gets less and less important to the overall Hubble evolution, but the photons are still there, getting more and more red shifted as time goes on. This gas of photons from the formation of atoms, a few hundred thousand years after the big bang, is the Cosmic Microwave Background Radiation (CMBR). A tiny fraction of these photons hit the earth and can be detected. Much of what we actually know about the early universe comes from studies of the CMBR. The first obvious thing to do is to measure the temperature, which turns out to be about 2.7°C, which is about 1000 times smaller than the temperature at which atoms come apart into ions. This means, since photon temperature and energy is inversely proportional to a, means that the universe today is about 1000 times bigger today than it was when atoms first formed. There is actually much more to this statement than meets the eye. It is an important prediction of the hot big bang model that CMBR looks like it has a temperature at all. The reason it does, even though the photons are no longer colliding very much, and are not in thermal equilibrium, is that the random distribution of photon energies that was present when the universe first became transparent is still there — just all the energies have been scaled down. This prediction has been confirmed by looking at the CMBR in many different regions of photon energy, and checking that the distribution of energies is what one would expect in a thermal distribution.

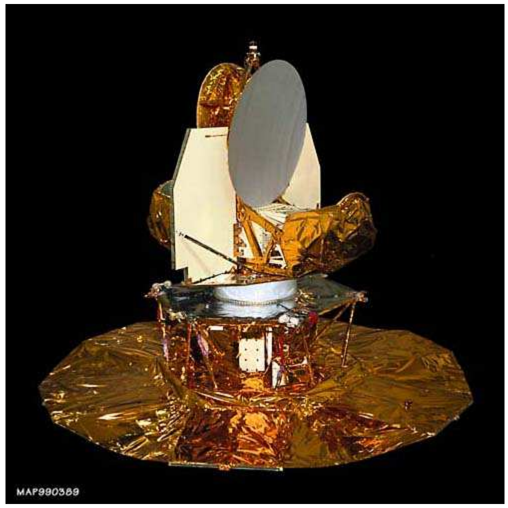
We cannot directly see the universe at scales smaller than 1/1000 the current scale. However, we can use the tools we have discussed to follow the universe back to small sizes and higher temperatures and energies. We can go back about a factor of a trillion (10^{12}) before we get to such high energies that have not directly seen the physics in laboratory experiments, so we are reasonably confident that we have the picture right back that far. And there are some observable consequences. For example, most of the nuclei of light elements (deuterium, helium, etc) were formed back when the universe was a million times smaller and protons and neutrons first started to stick together. The hot big bang picture predicts a specific pattern of abundances of these elements, which can be checked by looking for them out in the universe. This is much less direct and more problematic than direct observation, because one worries about what has happened to these nuclei in the 10 billion years since they were formed. Nevertheless, there is interesting information to be had in this way.[1]

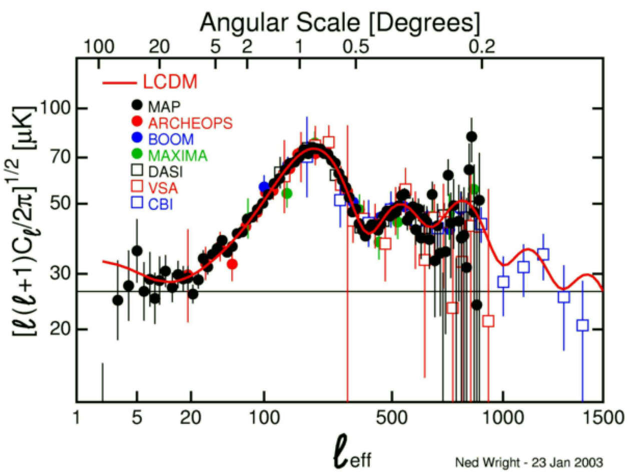
Another area of recent progress comes from the observation of small but regular temperature differences in the CMBR. These are associated with a funny kind of driven oscillation in the early universe. I think that what is going on can be described as follows.[2] The "oscillator" is the electron-proton-photon plasma. Before "recombination" when the electrons and protons (or heavier nuclei) form atoms, the electrons, protons and photons are tightly coupled together by collisions and form a gas which like air has a pressure (provided by the photons) and inertia (provided mostly by the photons but a little by the mass density the protons), and therefore a speed of sound. The "driver" is the primordial density fluctuations and gravitational instability of the dark matter. These fluctuations produce waves of density fluctuations in the plasma that propagate outward from a region of enhanced dark matter density like ripples on the surface of a pond disturbed by a rock. This wave propagation stops when the plasma recombines into atoms, gets frozen into the pattern of the CMBR on the sky, showing up as a wave in the temperature of the CMBR photons. These are known as acoustic peaks (check out the discussion and movie in this link). They are seen by

[1] Again, see Steven Weinberg's **The First Three Minutes** for more details.

[2] For a more detailed, but understandable description, see background.uchicago.edu/∼whu/intermediate/intermediate.html on the web or the Scientific American article by Hu and White, background.uchicago.edu/∼whu/Papers/HuWhi04.pdf.

instruments like the WMAP sattelite[3]





Eventually, we get back to such high temperatures that we really don't know how the physics works. But we suspect that one thing that shows up is interesting structure in the vacuum. We'll talk more about this in a moment.

Here is a web link to more information about the Cosmic Microwave Background:
http://background.uchicago.edu/
This is a very nice site maintained by a faculty member at Chicago.

Temperature and phases

The plasma that appears at high temperature in the early history of the big bang is an example of a "phase" of matter The idea of "phase" is one of the most familiar concepts in the physics of matter. You know from childhood about the three phases of H_2O - water-ice-steam. And you know that phases can change depending on temperature and pressure. This can have a big effect because physical properties can change dramatically in a "phase transition" as you go from one phase to another. I will close by showing you a couple of phase transitions.

[3]See http://www.astro.ucla.edu/∼wright/CMB-DT.html

lecture 26

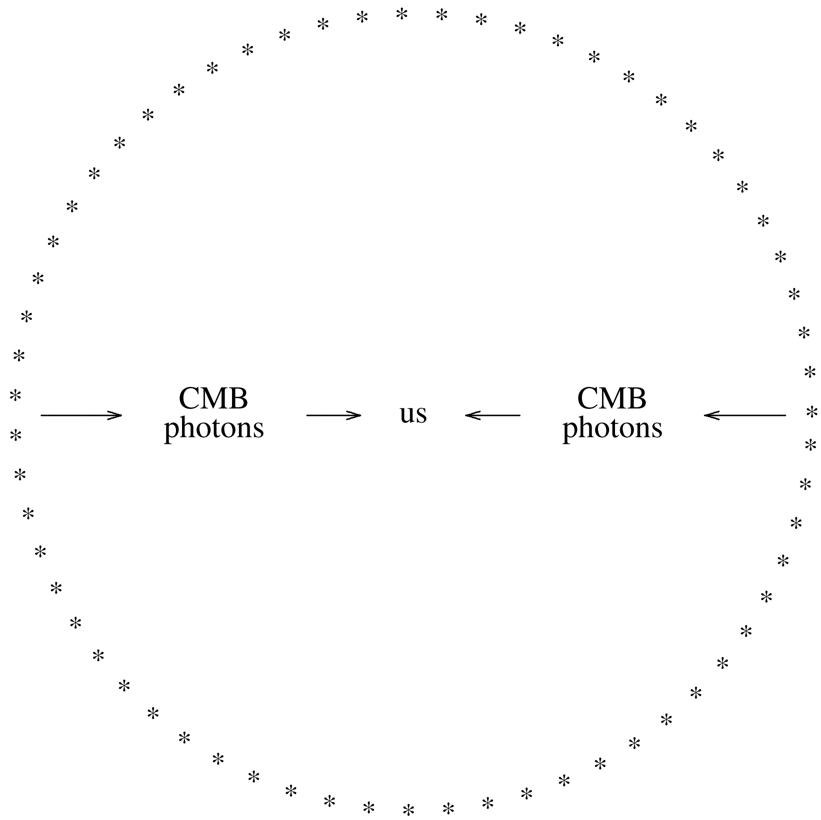
Topics:

- Before the big bang
- Inflation
- The cosmological principle and the Taylor expansion
- The cosmological constant
- The laws of physics and the Taylor expansion

Before the big bang

One reason that the cosmological principle has historically been treated as a sort of philosophical assumption is that until recently it was hard to imagine any sort of reasonable physics that could give rise to it. The reason that it seems difficult is special relativity. To see the difficulty, let us first think about the physics in a region of the universe that lies beyond the boundary of the observable universe. By definition, this is a region that we cannot see, because light from this region has not had time to reach us since the big bang. But according to special relativity, that means there is no way that any information can have been transferred from here to there or vice versa. This makes it hard to imagine any physical process that could establish the equivalence of this region of the universe with ours. Yet the cosmological principle requires that these two regions be equivalent. This does not mean that the cosmological principle is inconsistent. It just means that whatever established the cosmological principle did so as a kind of initial condition — before the big bang started the Hubble expansion. This makes the cosmological principle sound a bit fishy.

In fact, the situation is even worse. We don't really much care about whether the cosmological principle holds beyond the observable universe. But consider two regions on opposite sides of the observable universe.

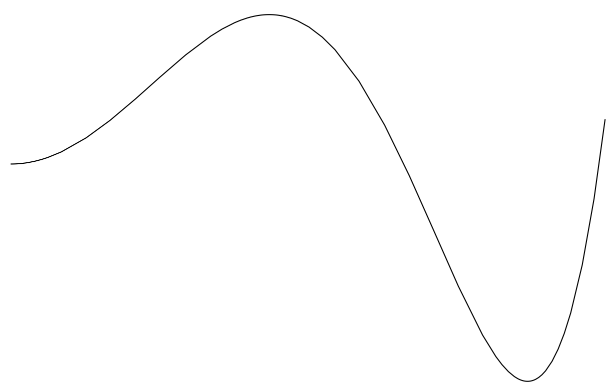


These are regions that we can see, and what we see is that the microwave background radiation from these two regions looks the same to very high accuracy. We can see directly that the cosmological principle holds for these two regions. But since light has just barely had enough time to reach us from these two regions, there is absolutely no way that light (and therefore information) can have gotten from one of these regions to the other in the history of the universe as we know it. Cosmologists refer this as the “Horizon Problem” because the boundary of the visible universe is a kind of horizon, beyond which we cannot see. The problem is that the similarity of the CMB from different regions of the cosmic horizon looks like a miracle, because no physics could have established it without violating the precepts of special relativity. The similarity must have been there as an initial condition before the big bang. Someone or something must have decreed that these regions looked the same before the conventional Hubble expansion started. What is going on here?

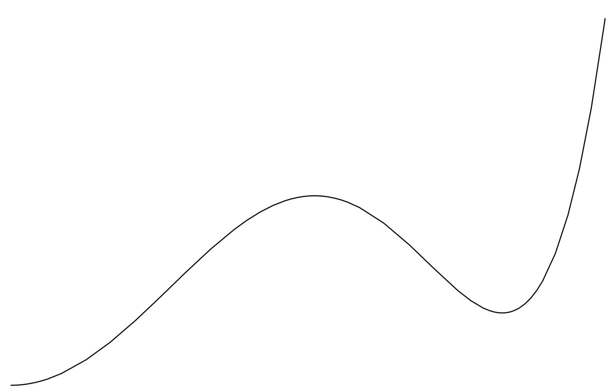
Inflation

One possible answer was suggested in the 80s by Alan Guth (now at MIT). I will describe the orginal version of Alan’s idea, even though it is not exactly the modern view of what happens because I think that it is the easiest to explain and is related to some interesting physics in the everyday world. Alan thought about the effect on cosmology of the funny structure of the relativistic, quantum mechanical vacuum. Like water, the vacuum is really a pretty complicated place. It has different phases. It breaks symmetries of the underlying interactions because it has structure that allows it to exist in different orientations in some odd quantum mechanical space. We have very strong reasons to believe that the vacuum can exist in different phases, each Lorentz invariant, but with different energy densities and other properties. At low temperatures, the phase we live in has the lowest energy and is stable (we hope). But at the high temperature of the big bang, one of the other phases might be the lowest energy state and be stable (like steam is more stable than liquid water at high temperatures). Such a state, if it exists, is called a "false vacuum."

Suppose that there is some continuous parameter (there may be many) that describes the configuration of vacuum. Then one way that multiple phases can happen is that energy of the vacuum as a function of this parameter has a couple of local minima, but that the curve depends on temperature. So at low temperatures it might look like



with the stable phase corresponding to the minimum on the right while at high temperatures it might look like



with the stable phase corresponding to the minimum on the left. This is the situation with water and steam in the real world. At high temperatures then right after the big bang, the vacuum state of the universe would typically be in the high-temperature "false" vacuum, and as the universe cools, there would come a time when the two vacuum states had the same energy, and after that, at lower temperatures, the true vacuum would be more stable. But in this situation, there is still an energy barrier between the two vacua so the transition from the false vacuum to the true vacuum is not immediate. The false vacuum is said to be "meta-stable." What happens then is that either as the result of a seed, or of random quantum mechanical fluctuations in the case of the universe, small regions of true vacuum can form and if they are large enough, they realize that they are more stable and grow and eventually permeate the whole space. Rain drops and the little bubbles that form when you boil water are examples of the same kind of phenomenon.

But one might think that even if the vacuum were in a different phase in the early stages of the big bang, that as the universe expands and cools, the usual vacuum would simply eventually reappear when things got cool enough. There is the interesting physics of metastability associated with the transition from one vacuum to another. But otherwise, the effect of the false vacuum should simply disappear when the universe cools. What Guth and others gradually realized is that this situation gets more complicated in the presence of gravity. It is true that eventually this false vacuum will decay into ordinary particles moving in the "true" low temperature vacuum with the minimum energy. But in the meantime, while the false vacuum fills space, something bizarre happens — the space full of false vacuum expands incredibly rapidly! Strange as this behavior is, we can actually understand it very simply from the equation for Hubble dynamics that we derived last time and used to discuss the evolution of the expansion of the universe. For a flat universe,[1] this can be written as follows:

$$\frac{\dot{a}}{a} = \sqrt{\frac{8\pi\, G\, \rho}{3}} \tag{1}$$

where a is any distance between points in space (distance between galaxies, whatever). We derived this assuming that ρ is the mass density of galaxies, but I assured you that it is true more generally if ρ (in units with $c = 1$) is interpreted as the energy density — basically because any kind of energy, not just rest energy (which is mass$\times c^2$) can produce and feel the force of gravity.

We have seen how this works if the energy density comes primarily from nonrelativistic matter or from radiation. Now lets consider the surprising effect of gravity on a universe in which the energy density is dominated by the energy of a "false vacuum." Here the difference is that as the universe expands, you just get more false vacuum, just as you get more true vacuum when the space in our present universe expands. But if the energy density of the false vacuum is the dominant contribution to ρ, that means that ρ does not fall off as a increases — it remains approximately constant, determined by the energy density of the false vacuum.

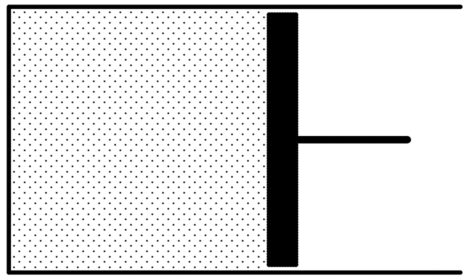
This sounds a little crazy. It is worth checking it with the relation we derived last time for the rate of change of the energy density,

$$\dot{\rho} = -3\frac{\dot{a}}{a}\,(\rho + p) \tag{2}$$

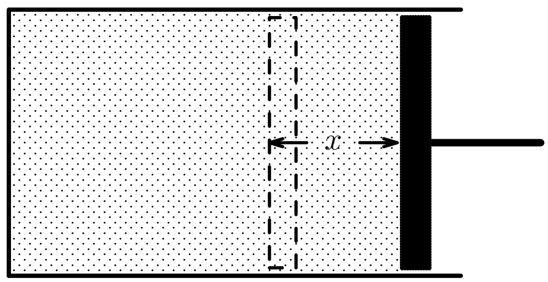
To check this, we need to know the relation between energy density and pressure in the false vacuum. Suppose that we have some false vacuum in a cylinder that we can expand with a piston

[1] We will discuss what happens if the universe is not flat shortly.

—



When we pull out the piston a distance x, we make more false vacuum — with volume xA where A is the area of the piston. But this costs energy $\rho x A$.



Thus the piston must be pulling back on us with a force ρA. **But that means that the pressure associated with the false vacuum is negative — equal to $-\rho$.** This weird result makes sense, because if we put $p = -\rho$ into (2), we find that $\dot{\rho} = 0$ — which is just what we expect — the energy density of the false vacuum is constant, independent of the volume of space.

So now that we are convinced that it makes a kind of weird physical sense to have ρ constant, we can go back to (1), which becomes

$$\frac{\dot{a}}{a} = \sqrt{\frac{8\pi\, G\, \rho}{3}} = \eta \tag{3}$$

where η is a constant! This is really easy to solve.

$$\dot{a} = \eta\, a \qquad \Rightarrow \qquad a \propto e^{\eta\, t} \tag{4}$$

The expansion of the space is actually exponential in time!!!! As you know, exponentials can get big quickly. This process is called “inflation” — during inflation, while the false vacuum lasts, space, and the universe along with it expands by some enormous factor.

Wait a minute — this is even crazier than the constant ρ. Isn’t gravity supposed to slow down the expansion of the universe? Let’s see by looking at the equation we derived last time for the second derivative of a:

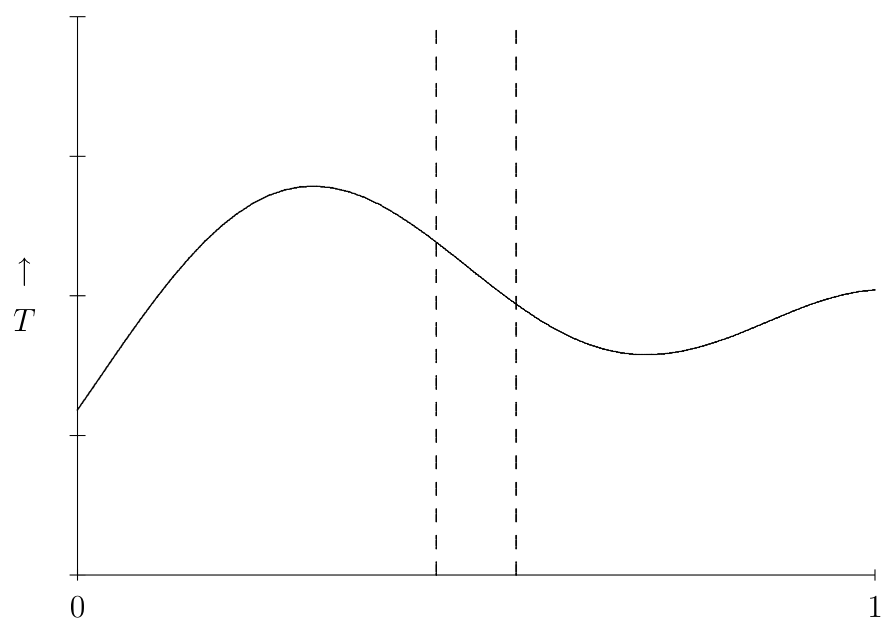
$$\frac{\ddot{a}}{a} = -\frac{4\pi\, G}{3}(\rho + 3p) \tag{5}$$

You see now that if we put in $p = -\rho$, the right hand side becomes positive (and you can check that it has the right value, η^2). This sign change is another effect of the negative pressure. This is the crazy physics of inflation. Positive pressure contributes to the slowing down of expansion, like energy density (only more so because of the factor of 3). But the false vacuum has NEGATIVE pressure, which gives a kind of antigravity effect! It’s weird, but all the equations are consistent.

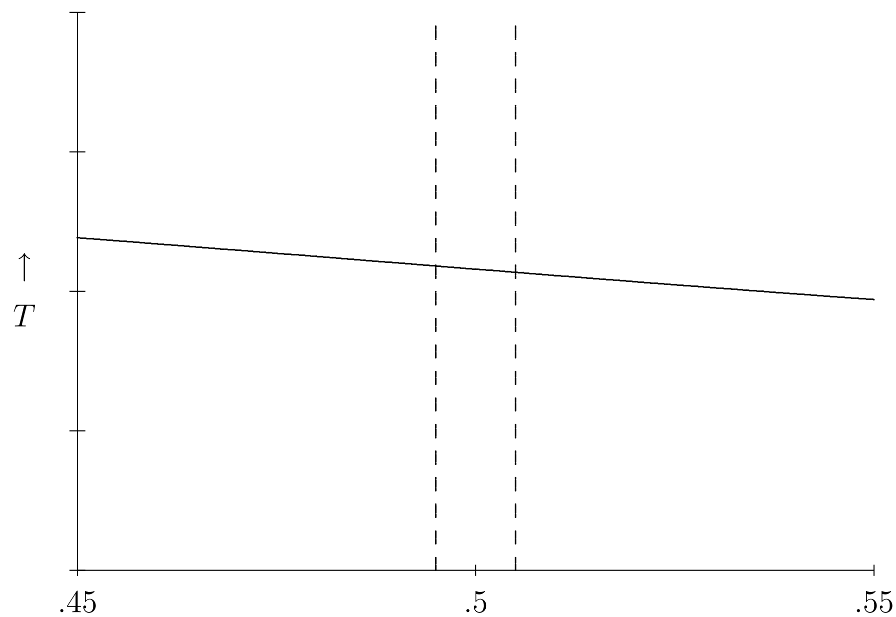
It doesn't really matter that we have used the flat-space ($C = 0$) version of the Hubble dynamics in (1). If the space is not flat at the beginning of an inflationary interlude, the rapid inflation will flatten it out, in the same way that the surface of a balloon gets flatter as we blow it up. Thus one consequence of inflation is the prediction that after inflation, when the false vacuum decays into normal vacuum and matter, the universe is very flat, so that the density must equal the critical density to a very good approximation. This is the theoretical prejudice in favor of flatness that I mentioned in the previous lectures.

The cosmological principle and the Taylor expansion

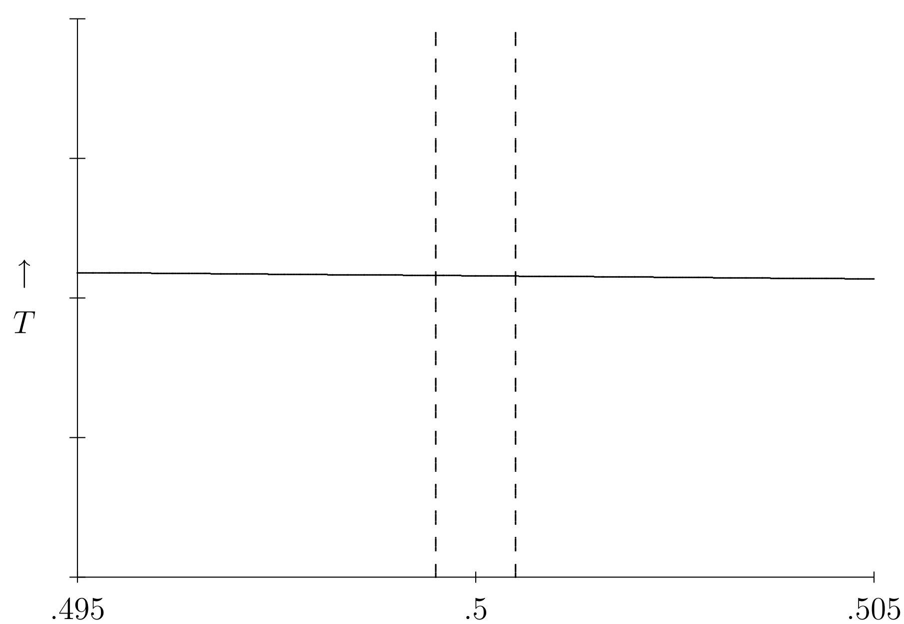
How does inflation help us with the Horizon Problem? Let's think about our two regions on opposite sides of the observable universe. In an inflationary universe, these two distant regions were actually very close together before inflation, so it is very reasonable that they are surprisingly similar. To put this another way, suppose that before inflation, the distance over which there was significant variation of the physical properties of the universe was R. During inflation, as the space expands, this distance scales up with the general expansion of the space. Thus the distance over which there is significant variation, no matter what it was originally, becomes huge. But that means that the derivative of any property of the universe (for example, the temperature, T) with respect to position becomes very small after inflation. Suppose that the temperature looks like this as a function of some position variable before inflation.



If the universe inflates by a factor of 10, then the little piece between the dotted lines gets blown up to the same size as the original region, so the temperature looks like this.



If the universe inflates by another factor of 10, this happens again.



And so on. I hope that this rings a bell from our discussion long ago of the harmonic oscillator. This is the Taylor expansion in action. The temperature, and any other important quantity, if it is described by a function $\tilde{T}(\vec{r})$ before inflation and doesn't change during inflation of the space by a factor of κ except due to the expansion of the space, then after inflation

$$T(\vec{r}) = \tilde{T}(\vec{r}/\kappa) \tag{6}$$

So that each derivative brings down a factor of $1/\kappa$. Because the derivatives are tiny after inflation, the function can be approximated by the first constant term in a Taylor expansion.

$$T(\vec{r}) = T(0) + \cdots \tag{7}$$

Inflation makes all the higher terms tiny. As long as there is only a very tiny variation of these quantities over the size of the visible universe, the visible universe looks to us as if it satisfies the cosmological principle. In the inflationary universe, it is the exponential expansion and the Taylor expansion that establishes the very smooth initial conditions of the big bang.

I think that inflation probably has something to do with very early history of the observable universe but that there are a lot of pieces of the puzzle still missing. There are many confusing issues involved in understanding how the inflationary era turns into the conventional big bang. It is a pretty wild collection with names like "eternal chaotic inflation." The idea is that the vacuum energy fluctuates because of quantum mechanics, and if, in some region, the vacuum energy is large enough, the negative gravity induced by its negative pressure causes it to inflate and grow wildly. Because the inflating regions are growing so fast, at any given time most of the volume of the universe is inflating. But in a few places, the vacuum energy settles down to its minimum value and something like our universe results. Absolutely nuts! Suffice it to say that some of these issues are actually being tested by the precise satellite studies of the cosmic microwave background. But many others will likely remain mysterious. This is a creation myth after all. We shouldn't expect to understand everything!

I do think that the general picture of an inflated universe is interesting — even liberating. Think for a minute about what this means. We have gotten used to thinking about our universe as unimaginably big. But in the inflationary view, our visible universe is actually incredibly tiny. All the billions of galaxies whose light has reached us since the beginning come from what was initially a tiny spot in a much bigger system, and beyond the edge of our visible universe, the universe goes on and on — whether it is infinite or not, it is definitely exponentially larger than what we can see. The cosmological principle is no longer needed — instead an approximate uniformity arises because of inflation. Every point in our visible universe is equivalent to every other because in the cosmic scheme of things, they were once nearly the same point! But we should not assume that distant parts of this inflated universe are similar to the little patch we live in. They may be very different indeed. Who knows whether even the laws of physics itself are constant over the absurdly large size of an inflated universe. Maybe if we could look sufficiently far away in various directions, we would find an infinite variety of completely different laws. Some of these might give rise to interesting but very different worlds. There is something about this picture that I find appealing — but since all these issues are invisible inside the confines of the visible universe, these questions are not very scientific. We will come back to this in a few minutes.

The cosmological constant

There is one more possibly very important thing to say about the energy density of the vacuum. In the discussion above, we have implicitly assumed that the true vacuum, after it settles down to its equilibrium value, has zero energy density. But why should this be true? Why should the true vacuum, the true lowest energy state of the universe, the state with absolutely nothing in it — why should this vacuum have zero energy density. This question is only meaningful in the presence of gravity. In Newtonian mechanics, as you know, only differences in energy make a difference. We can always redefine the energy by subtracting a constant from everything. But as we have seen with the false vacuum, in the presence of gravity, this is not true. The energy density and the negative pressure associated with it have a gravitational effect. And no one has ever been able to come up with a convincing reason why such an energy density should not be there. In fact, it was suggested by Einstein himself, early in the history of general relativity. It is called the 'cosmological constant."

Einstein originally suggested the cosmological constant for what turned out to be the wrong reason. He felt, for philosophical reasons, that the universe should not change with time. He didn't know about the Hubble expansion, which was discovered later. But with ordinary gravity, there is no way to make a universe that stands still. If it starts at rest, gravity will cause it to contract. But as we have seen, an energy density of the vacuum produces a repulsive gravitational effect because of its negative pressure. This can, if properly chosen, exactly cancel the effect of gravity, allowing for a stationary universe.

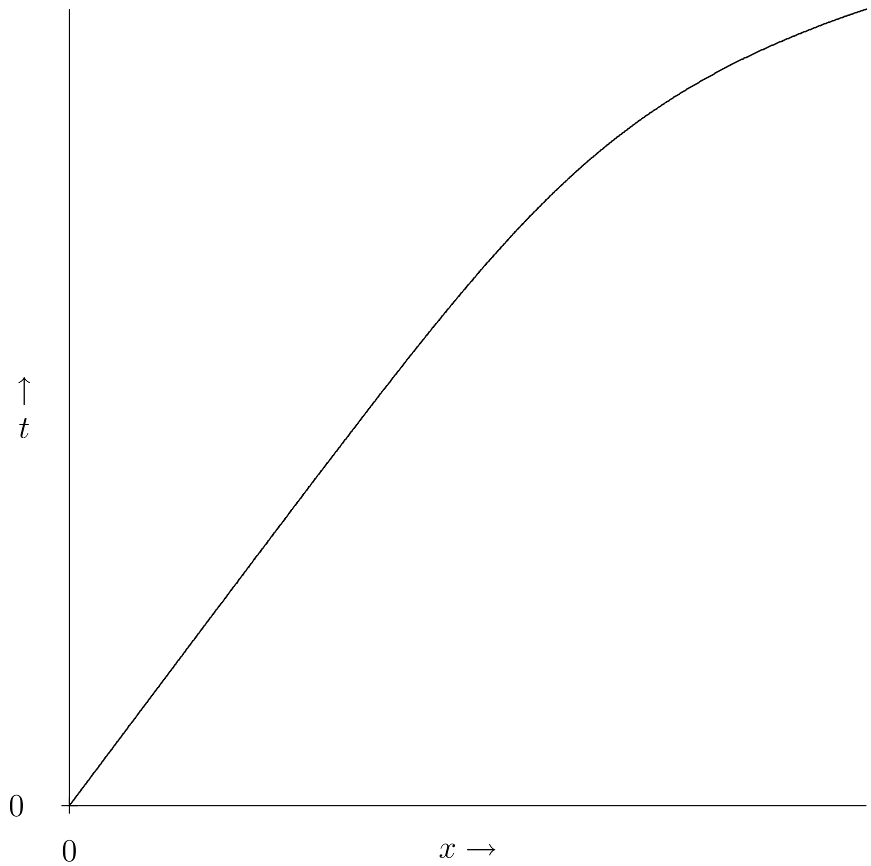
Of course this original motivation for the cosmological constant has long since gone away. It is sometimes referred to as "Einstein's mistake" because it seems very ugly to theorists like me. I'll come back below to why it seems so ugly. Nevertheless, a decade ago, astronomers published evidence that it might not be a mistake after all. By studying type Ia supernovas[2] in distant galaxies, astronomers found evidence that the Hubble expansion is not slowing down as one would expect if the energy density of the universe is dominated by any kind of matter. Rather, it looks as if the expansion is speeding up!

One of the the things that astronomers actually see is that the brightness of distant supernovas with red shifts between 0.3 and 1 is lower that what would be expected by computing their distance using their red shift. This is interpreted to to mean billions of years ago when this light was emitted, the relation between red shift and distance was different from what we would expect from extrapolating back from what we see now. For a given red shift, the objects were further away then than we would expect on the basis of the Hubble constant determined by looking at nearer (and thus more recently viewed) objects. Since we have to to farther out to get the same red shift, that means that the universe was expanding more slowly then. But that means that the expansion of the universe has been speeding up since the light from these supernovas was emitted, billions of years ago.

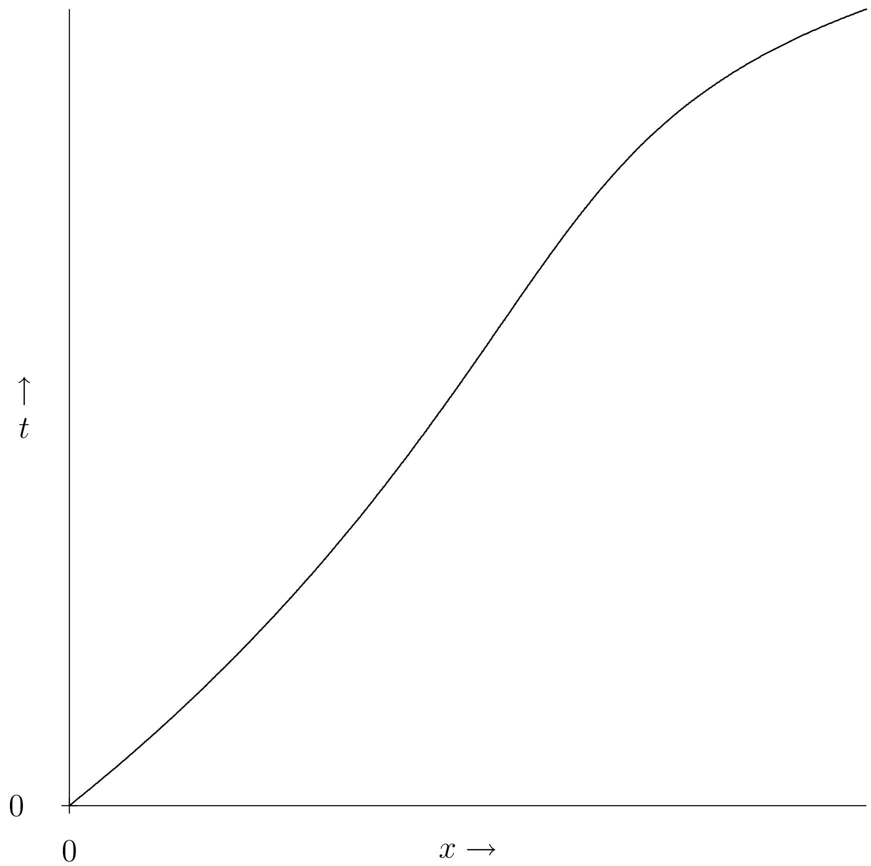
A picture may help. If the rate of Hubble expansion were constant, we could plot the position of a distant galaxy as a function of time and get a straight line. But if the expansion of the universe

[2]These are supernovas that occur when a white dwarf star gobbles up matter and increases in mass beyond the point where the pressure from atomic matter that keeps the star from collapsing can compete with the gravity that is trying to squeeze it. When gravity finally wins, the result is a huge explosion that at least theoretically should always produce about the same amount of energy and should therefore provide a "standard candle" that allows astronomers to determine its distance by measuring its brightness.

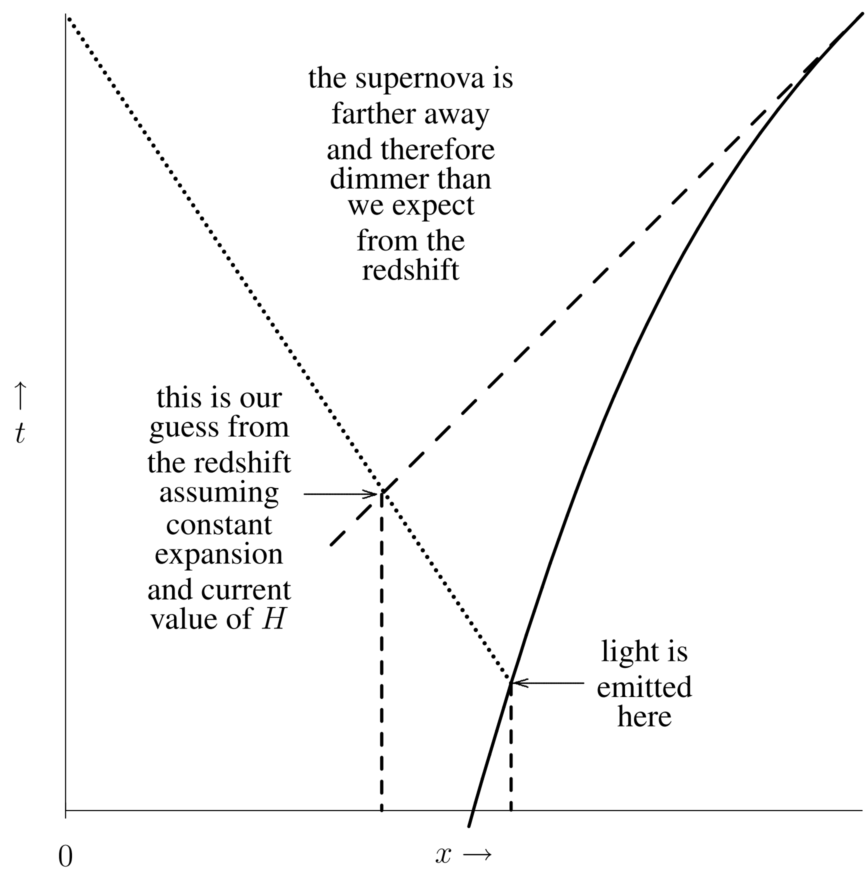
is accelerating, we get something like this (with time going up):



While this is going on, we are at the origin, just going up along the vertical axis. In fact, we expect the expansion of the universe to decelerate initially, when the ordinary energy dominates, but eventually to accelerate when the ordinary energy density falls below the constant dark energy of the vacuum, so the picture might look something like this:



Let's not worry about this complication for the moment, and just focus on just the last part of this trajectory, where things are accelerating. I hope that the picture will make it clear what is going on in this experiment. The dotted line represents the light ray reaching us from the distant supernova.



Two groups have independently seen this effect. One is a local group (involving our department chair Chris Stubbs): http://arxiv.org/abs/astro-ph/9805201. The other is a group from Berkeley: http://arXiv.org/abs/astro-ph/9812133.

If all this is right, it implies that about 70% of the energy density of the universe, which it seems is roughly equal to the critical density required for flatness, comes from the energy of the vacuum. This is a bizarre, crazy result that I have a lot of trouble believing, for reasons that I will come back to. However, it has been around for almost a decade now as the number of supernovas on which the estimate is based has grown from a few to hundreds. There is also now considerable indirect support for this view from the consistency of the measured anisotropy of the CMBR and the gravitational lensing studies of dark matter with a model in which 30% of the energy density of the universe is in the form of cold dark matter and the rest is energy density.

It is still hard to believe, but who knows.

The laws of physics and the Taylor expansion

Now finally, I want to return to the question that we discussed at the very beginning of the course. Why is $\vec{F} = d\vec{p}/dt$? Why is Newton's law a formula for acceleration? Of course, we know that Newton's law is not right. It ignores special relativity and quantum mechanics, for example. But in fact, our current theories of relativistic quantum mechanics are based on Lagrangians that are really rather straightforward generalizations of those that we use to derive Newton's law. The Lagrangians that we use to describe the world still depend just on coordinates and their first time derivatives. The equations of motion are thus still equations for "acceleration." So in a certain sense, Newton has survived the revolutions of special relativity and quantum mechanics, and the question still remains an interesting one. I certainly don't know the answer to this question. But I think that it is related to a much deeper question, which may be the central mystery about the way the universe works.

This mystery takes a bit of explaining. We have already talked about the two constants, $\hbar$ and c, that are built into the way the world works. We are used to setting $c = 1$. But nature is also telling us to set $\hbar = 1$. If we adopt these sensible units (which are sometimes called "particle physics units" - you may have read about them in Morin's book), then all dimensional quantities can be related. For example, we can express everything in terms of mass. The properties of our world are primarily determined by the masses of the electron and the proton, and a few numbers, like the fine structure constant, $\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$. Almost all of the physics of the everyday world involves combinations of these basic parameters. But one thing that is different is gravity. Gravity is described by the gravitational constant G, which in particle physics units is proportional to $1/m_{\rm Planck}^2$ where the Planck mass is enormous, about 10^{19} times the mass of the proton. G, is very very very tiny compared to any quantity with the same units that we might construct out of the parameter that describe the rest of our everyday world.

Why should we care about gravity? Well, aside from the fact that it keeps us from flying off into space, there is a theoretical problem associated with gravity. It seems to be impossible to put special relativity, quantum mechanics, and gravity together consistently, without changing the rules in some way. This suggests that the Planck mass is the basic scale at which really interesting new physics, some change of the rules beyond special relativity and quantum mechanics, is happening. This is a very dicey argument, for various reasons,[3] but let's assume that it is right. Then special relativity and quantum mechanics, and with them Newton's law, are just approximations that are true for masses much smaller than the Planck mass.

Now the important point is this. Every dot in the Lagrangian, every derivative with respect to time, in particle physics units, has units of mass (because a derivative is one over a distance). But if the fundamental scale is the Planck mass, each dot should come generically with a power of $1/m_{\rm Planck}$. Because $m_{\rm Planck}$ is so huge, the terms with more than the minimum number of dots can be ignored. They are there, but their effects are very small, and we don't see them.

If this is right, then $\vec{F} = d\vec{p}/dt$ is just an approximation, but it is a very good one because gravity is so weak. The central mystery, then, is why is gravity so weak? Why is the Planck mass so very much larger than all the other masses that we care about in physics? This seems crazy. Where does the tiny dimensionless ratio of the proton mass to the Planck mass, about 10^{-19} come from? Who know? Maybe one of you will figure it out, or maybe you will show that it is not the

[3]For example, maybe someone will discover a clever way of doing it. People are trying. Maybe the string theorists have already done it, but they have really changed the rules, so this would just be an example of what I am saying. Or maybe the rules change again well before the Planck scale, and gravity emerges in some complicated way from physics at smaller masses. There are also some interesting ideas of this sort on the market today.

right question.

In closing, let me finally come back once more to the cosmological constant. If the recent reports of a nonzero energy density of the vacuum are correct, this is ever crazier that the mystery of the proton mass. To see this, lets convert this cosmological constant to particle physics units. The statement is that the cosmological constant is about 70% of the critical density. The critical density is given by

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{1}{T^2} = \frac{8\pi G\rho_c}{3} \tag{8}$$

where T is the Hubble time, very roughly 10^{10} yr. So

$$\rho_c \approx \frac{3}{8\pi G T^2} \tag{9}$$

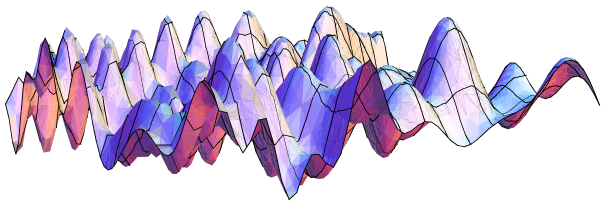
The $1/G$ is just m^2_{Planck}. One year is about $\pi \times 10^7$ seconds. The way I remember how to do these unit conversions is to convert things first to a funny unit of mass - billion electon volts - GeV - which is about the mass of the proton. In these units, one second is about 10^{24} GeV^{-1}. But the Planck mass is even larger, about 10^{19} GeV, so 1 second is about 10^{43} m^{-1}_{Planck}. Thus a year is about 10^{50} m^{-1}_{Planck} (we can certainly drop the factor of π) and the Hubble time is about 10^{60} m^{-1}_{Planck}. Thus

$$\rho_c \approx 10^{-120}\ m^4_{\text{Planck}} \tag{10}$$

This is an even smaller dimensional scale than the proton mass. That is why it seems so crazy. Where does this tiny number come from? This situation is much worse even than the ratio of 10^{19} between the Planck mass and proton mass. If we had to guess from first principles the energy density of the universe, we would guess that the scale is set by the Planck mass, so we would guess m^4_{Planck}. We missed it by about a factor of 10^{120}. This wasn't a very good guess! I think it is probably the worst guess in the history of mathematical science. What went wrong? Nobody knows. But it is suspcious that no only is it very small, but it is not so different from the critical density of the universe today.

There are two kinds of ideas that people have discussed to try to explain away this puzzle. Many theorists believe that it has something to do with an odd kind of symmetry - called supersymmetry - which would imply an equivalence between matter particles and force particles. If this symmetry were exact, the cosmological constant would be zero. Unfortunately, we don't see this symmetry - it is certainly broken in our world - and nobody have found a plausible explanation of how the symmetry could be broken, and still force the cosmological constant to be small.

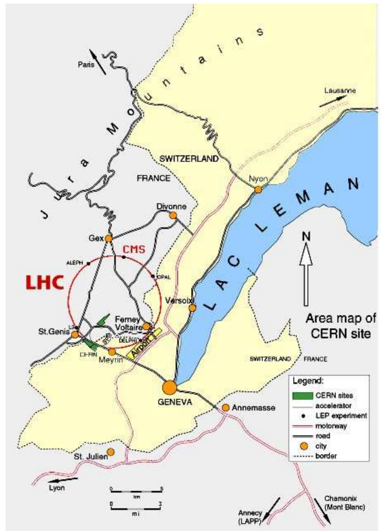
The other set of ideas is even crazier and goes by the term "Anthropic argument." This is an updated version of an old idea that the laws of physics are what they are because if they were different, we wouldn't be here to study them. The modern version of this goes something like this. Suppose that vacuum is a much more complicated thing than I have described to you, and actually the structure of the vacuum in which we live depends on many many parameters, and the energy of the vacuum is a complicated "landscape" with lots of peaks and basins. A two dimensional verison really does look like a landscape —



but of course the multidimensional version is much more complicated and hard to visualize. If the vacuum looks like this, and quantum mechanics is causing fluctuations in this complicated structure, then most regions in the huge chaotic universe are somewhere up on a hill, inflating away like mad because of negative gravity. but in some regions the universe settles down into one of the basins, producing matter and radiation which dominates the energy density. Then this region stops inflating, at least for a while, until the ordinary Hubble expansion of this region of the universe dilutes the matter and radiation and the cosmological constant dominates the energy density. Then the expansion accelerates at a rate determined by the value of the energy density at the bottom of the basin. This happens a huge number of times in different regions of the universe, in a huge number of different basins, at random. Each basin will have a different of the cosmological constant. This kind of universe is a kind of huge experimental laboratory with different experiments going on each of the regions where the universe settles down. One constraint that many people think is relevant has to do with the formation of galaxies. If the cosmological constant is too big in a basin, then the space continues to grow so fast that matter doesn't have a chance to fall into separate galaxies. Then stars don't form, planets don't develop, and nobody gets to teach Physics 16. If you believe that this is necessary, it goes a ways towards explaining why the cosmological constant is so small.

I am not sure that this is science. But one could imagine getting enough information about structure of this landscape to make predictions rather than just qualitatively explaining what we see. That would be interesting.

At any rate, it seems clear that we need to understand the vacuum much better. If we are really lucky, we will get some hints to what is going on by seeing some surprising property of the vacuum at the LHC in the next few years.



Above are a map and an arial photo of the CERN site, near the Geneva airport, where the LHC is being constructed. This is a monster project involving many thousand scientists and engineers.

Below is a photo in the tunnel - most of the 10 miles of it looks like this.

It is pure speculation at this point to think that the LHC may teach us something about the dark energy. But it could happen. Certainly, the LHC will be our first opportunity to directly probe the mysterious physics of the vacuum, of which the dark energy seems to be another manifestation. That should make the next ten years an incredibly exciting time for fundamental physics.