

MITOCW | 6. The principle of equivalence.

[SQUEAKING]

[RUSTLING]

[CLICKING]

SCOTT

So we're just picking up where we stopped last time. So we are beginning to discuss

HUGHES:

how we are going to sort of do a geometrical approach to physics, using a more general set of coordinates now. So we began talking about how things change when I discuss special relativity, so for the moment keeping ourselves just at special relativity.

We, by the way, are going to begin lifting our assumptions that it is simply special relativity fairly soon. But to set that up, I need to start thinking about how to work in more general coordinate systems. So we're going to do it in the simplest possible curvilinear coordinates. So it's basically just going from Cartesian coordinates in the spatial sector to plane polar coordinates.

One of the things which I have emphasized a few times, and I'm going to continue to hammer on, is that these are a little bit different from the curvilinear coordinates that you are used to in your past life. In particular, if I write out the displacement, the little vector of the displacement element in the usual way, I am using what's called a "coordinate basis," which means that the vector dx is related to the displacement, the differential of the coordinates, by just that thing contracted with all the basis vectors.

And so what that means is I have a little displacement in time, which looks normal. Displacement in radius, which looks normal. Displacement in the z direction, which looks normal, and a displacement in an angle, which does not. In order for this whole thing to be dimensionally consistent, that's telling me e^ϕ has to have the dimensions of length. And that is a feature, not a bug.

Last time, we introduced the matrix that allows me to convert between one coordinate system and another, so just basically the matrix-- it's sort of a Jacobi matrix. It's a matrix of partials between the two coordinate systems. And this idea

that things are-- they look a little weird.

So the way I did that was I didn't actually write it out, but I did the usual mapping between x , y and r and ϕ , worked out all of my derivatives. And sure enough, you've got something that looks very standard, with the possible exception of these r 's that are appearing in here. So notice the elements of this matrix. These do not have consistent units-- again, feature, not bug.

This guy is basically just the inverse of that. This is the matrix that affects the coordinates in the opposite direction. And notice in this case, you have some elements where their units are 1 over length.

So let's just continue to sort of catalog what some of the things we are going to be working with look like in this new coordinate representation. And this will lead us to introduce one of the mathematical objects that we are going to use extensively as we move forward in studying this subject. So what I want to do next is look at what my basis vectors look like.

So what I want to do is characterize what my e_r and my e_ϕ look like. And these are going to look very familiar from your intuition, from having studied things like E&M in non-Cartesian coordinates. So your e_r is just related to the original Cartesian basis vectors, like so.

And if you like, you can easily read that out by performing the following matrix multiplication on the original Cartesian basis vectors. Your e_ϕ perhaps looks a little wacky. So you can see the length coming into play there.

A good way to think about this is if your intuition about basis factors-- I have to be careful with this language myself-- your intuition about basis vectors is typically that they are unit vectors. These are not unit vectors. They do form a nice basis, but they are not unit vectors.

In particular, the basic idea we're going to go with here is that e_ϕ , it's always going to sort of point in the tangential direction. But no matter where I put it in radius, I want that vector to always sort of subtend the same amount of angle. In order to do that, its length needs to grow with r . So that's where that's a little bit different from your intuition. And there's a very good reason for this, which we will

get to, hopefully, well before the end of today's class.

So last time, when we first began to talk about tensors a couple of lectures ago, the first tensor I gave you-- so confining ourselves just to Cartesian coordinates-- was the metric, which was originally introduced as this mathematical object that came out of looking at dot products between basis vectors. It's essentially a tensor that allows me to feed in two displacements and get the invariant interval between those displacements that comes out of that.

I am going to continue to call the dot product of two basis vectors the "metric." But I'm going to use a slightly different symbol for this. I'm going to call this $g_{\alpha\beta}$.

In the coordinate representation that we are using right now, so in plane polar coordinates, this becomes-- you can work it out from what I've written down right here. This is just the diagonal of minus 1, 1 r^2 1. So this equals dot here. This is-- I'll put PPC for plane polar coordinates under that.

And then using that, you know that you can always work out the invariant displacement between two events. It's always going to be the metric contracted with the differential displacement element. And this is going to be $-dt^2 + dr^2 + r^2 d\phi^2 + dz^2$.

That, I hope, makes a lot of sense. This is exactly what you'd expect if I have two events that are separated in plane polar coordinates by dt , dr , $d\phi$, dz . This is what the distance between them should be.

So the fact that my basis vectors have this slightly annoying form associated with them, it all sort of comes out in the wash here. Remember at the end of the day, if we think about quantities that are representation independent-- and that's going to be the key thing. When you assemble scalars out of these things, the individual tensor components, they can be a little bit confusing sometimes.

They are not things that we measure. They are not things that really characterize what we are going to be working with. And we really want to get into the physics, unless we're very careful about it. This is something you can measure. And so sure enough, it comes out, and it's got a good meaning to it.

Let me just wrap up one last thing before I talk about sort of where we're going with this. So just for completeness, let me write down the basis one forms. Just as the basis vectors had a bit of a funny form associated with them, you're going to find the basis one forms likewise have a bit of a funny form associated with them.

And the way I'm going to get these-- and so these are going to be the Cartesian basis one forms-- basically, I'm not carefully proving all these relations at this point, because you all know how to do that. I'm just using line up the indices rule. And when you do that, you get this. And likewise, your basis one form for the axial direction, I'll just write down the result. It's going to look like this.

So the key place where all of this-- so right now, these are all just sort of definitions. Nothing I've done here should be anything that even approaches a surprise, I hope, just given the you guys have done-- the key thing that's probably new is all this garbage associated with coordinate bases, this extra factors of r and $1/r$ that are popping up. But provided you're willing to sort of swallow your discomfort and go through the motions, these are not difficult calculations.

The key place where all of this really matters is going to be when we calculate derivatives of things. It'll turn out there is an important rule when we talk about integrals as well a little bit later, but let's just say that. So for now, we'll focus on derivatives.

So all the derivatives that we've been looking at so far, we have, indeed, done a couple of calculations where we've computed the derivatives of various vector valued and tensor valued quantities. And it was helped by the fact that all the bases, when I work in Cartesian coordinates, are constant.

Well, that's not the case now. So now, we need to account for the fact that the bases all vary with our coordinates. So let me just quickly make a catalog of all the non-trivial-- there's basically four. In this one, where I'm just doing plane polar coordinates, there are four non-trivial derivatives we need to worry about.

One of them actually turns out to be 0. So the radial derivative of the radial unit vector is 0. But the ϕ derivative of the ϕ unit vector is not. you go and take the ϕ derivative of this guy, and you basically get-- take the ϕ derivative of this,

you're going to get this back, modulo factor of radius.

So I can write $d e_r d \phi$ as e_ϕ over r . If I take the derivative of e_ϕ with respect to r , I get e_ϕ back, divided by r . So the simplest way to write this is like so.

And finally, if I take the ϕ derivative of the ϕ unit vector, I get e_r back, with an extra factor of r thrown in. And a minus sign. So we're going to see a way of doing this that's a little bit more systematic later, but I want to just keep the simple example, where you can just basically by hand calculate all the non-trivial derivatives easily.

Of course, there's also a t unit vector and a z unit vector. But they're constants, so I'm not going to bother writing them out. All the derivatives associated with them are equal to 0.

So let's imagine now that I have assembled some vector. So I have some vector field that lives in this spacetime. And I'm using this basis. And so I would write this vector with components v_α . And let's let the-- so this is going to be a curvilinear coordinate system, so this will be plane polar coordinates being used here, plane polar coordinate basis vectors.

And what I would like to do is assemble the tensor that you can think of essentially as the gradient of this vector. So let's begin by doing this in a sort of abstract notation. So the gradient of this guy-- this is sort of ugly notation, but live with it.

Following what we have been doing all along, what you would want to do is just take the root of this whole thing. It's going to have a downstairs component on it. So attach to it the basis one form. If you prefer, you can write it using the d notation like I have there, but I just want to stick with the form I wrote in my notes.

Looking at this the way I've sort of got this right now, I can think of, if I don't include the basis one forms here, this should be the components of a one form. So this should be a kind of object. So let's just expand out that derivative.

Let's write it like this. So you just-- we haven't changed calculus. So when I do this, I'm going to basically use the old-fashioned Leibniz rule for expanding the derivative product of two things.

Here's the key thing which I want to emphasize-- in order for this whole thing to be-- for this to be a tensorial object, something that I couple to this basis one form, the sum of these two objects must obey the rules for transforming tensors. But the two objects individually will not. So this is an important point which I'm going to emphasize in slightly different words in just a few moments again. This is one of the key things I want you to get out of this lecture, is that when I'm taking derivatives of things like this, you've got to be a little bit careful about what you consider to be components of tensors and what is not.

Now as written like that, this is kind of annoying. So my first object has a nice basis vector attached to it. My second object involves a derivative of the basis vector.

However, something we saw over here is that derivatives of basis vectors are themselves proportional to basis vectors. So what I'm going to do is introduce a bit of notation. So let me switch notation slightly here.

So the beta derivative of e alpha can be written as-- in general, it can be written as a linear combination of basis vectors. So what we're going to do is define d-- I want to make sure my Greek letters are legible to everyone in the room here. So let me write this nice and clearly.

$d_{\beta} e_{\alpha}$, I'm going to write that as capital gamma mu beta alpha e mu. This gamma that I've just introduced here in this context is known as the Christoffel symbol. Fact I'm calling this a symbol, it's got three indices on it.

You might look at it and go, ooh, smells like a tensor. Be a little bit careful. In much the same way that these two terms are not individually components of a tensor, but their sum is, this guy individually is actually not a component of a tensor, but when combined with other things, it allows us to assemble tensors.

So for our plane polar coordinates, there are exactly three non-zero Christoffel symbols. So $\gamma_{\phi r \phi}$ is equal to $1/r$, which is also equal to $\gamma_{\phi r \phi}$. $\gamma_{r \phi \phi}$ is minus r . And you can basically just read that out of that table that I wrote down over there. All the others will be equal to 0.

Now from this example, this is what it makes it smell like every time you introduce a new coordinate representation. You're going to need to sit down for an hour and a

half, or something like that, and just work out all the bloody derivatives, and then go, oh, crap, and read out all the different components of this thing, and assemble them together. There actually is an algorithm that we will get to at the end of this class that allows you to easily extract the Christoffel symbols provided you know the metric.

But right now, I just want to illustrate this thing conceptually. The key thing which you should know about it is that it is essentially the-- I almost said the word "matrix," but it's got three indices. It's a table of functions that allows me to relate derivatives of basis vectors to the basis vectors. So before I go on to talk about some of that stuff, let's take a look at the derivative a little bit more carefully.

So the derivative of the vector-- so let's basically take what I've written out up there. I'm going to write this as the beta derivative of vector v . And I can write that as the beta derivative of e of v alpha-- so the first term where the derivative hits the vector components.

And then I've got a second term where the derivative hits the basis. I'm going to write this like so. This is sort of annoying. One term is proportional to e alpha, one is proportional to e mu.

But notice, especially in the second term, both alpha and mu are dummy indices, so I'm free to relabel them. So what I'm going to do is relabel alpha and mu by exchanging them. As long as I do that consistently, that is totally kosher. And when I do that, I can factor out an overall factor of the basis object.

This combination that pops up here-- so we give this a name. And this is a combination which, by the time you have finished this semester, if you don't have at least one nightmare in which this name appears, I will not have done my job properly. This shows up a lot at this point.

This is called the "covariant derivative." And it shows up enough that we introduce a whole new notation for the derivative to take it into account. I'm going to call this combination of the partial derivative of v and v coupled to the Christoffel symbol-- I'm going to write this using the, if you're talking LaTeX, this would be the nabla operator.

So I made a point earlier when we were talking about derivatives a couple of weeks ago that we were reserving the gradient symbol for a special purpose later. Here it is. So whenever I make a derivative that involves the gradient symbol like this, it is this covariant derivative.

And the covariant derivative acting on vector components, it generates tensor components. Partial derivative does not. And what I'm going to do, just in the interest of time-- it's one of those calculations that's straightforward but fairly tedious-- I have a set of notes that I meant to make live before I headed over here, but I forgot. I have a set of notes that I'm going to show on the website by this evening which explicitly works out what happens when you apply the coordinate transformation using that-- it's been erased-- when you use that L matrix to construct the coordinate transformation between two representations.

If you try to do it to partial derivatives of vector components, basically what you find is that there's an extra term that spoils your ability to call that-- it spoils the tensor transformation law, spoils your ability to call that a tensor component. So the partial on its own doesn't let you. You get some extra terms that come along and mess it all up.

On next p set, you guys are going to show that if you then try to apply the tensor transformation law to the Christoffel symbols, you get something that looks tensorial, but with an extra term that spoils your ability to call it tensorial. There's a little bit of extra junk there.

But the two terms exactly conspire to cancel each other out so that the sum is tensorial. So part one of this will be notes that I post to the website no later than this evening. Part two, you guys will do on the p set.

So just saying in math what I just said in words, if I do this, like I said, you basically will eventually reach the point where what I am writing out right now will become so automatic it will haunt your dreams. Wait a minute, I screwed that up. It's so automatic I can't even write it properly.

Anyhow, something like that will-- modulo my typo-- that should become automatic. And the key thing which I want to note is that if I take these guys, and I attach the appropriate basis objects to them, this is an honest-to-god tensor. And so this

derivative is itself an honest-to-god tensor.

A typical application of this, so one that will come up a fair bit, is how do you compute a spacetime divergence in each coordinate system? So suppose I take the divergence of some vector field v . So you're going to have four terms that are just the usual, like when you guys learned how to do divergence in freshman E&M in Cartesian coordinates. You get one term that's just $dv_x dx$, $dv_y dy$, et cetera.

So you got one term that looks just like that, and you're going to have something that brings in all of Christoffel symbols. Notice the symmetry that we have on this one. Actually, there is Einstein summation convention being imposed here.

But when we look at this, there's actually only one Christoffel symbol that has repeated indices in that first position. So when I put all this together, you wind up with something that looks like this. So go back and check your copy of Jackson, or Purcell, or Griffith, whatever your favorite E&M textbook is.

And you'll see when you work in cylindrical coordinates, you indeed find that there's a correction to the radial term that involves $1/r$. That's popped out exactly like you think it should. You have a bit of a wacky looking thing with your ϕ component, of course.

And let me just spend a second or two making sure. It's often, especially while we're developing intuition about working in a coordinate basis, it's not a bad idea to do a little sanity check. So here's a sanity check that I would do with this.

If I take the divergence, I take a derivative of a vector field, the final object that comes out of that should have the dimensions of that vector divided by length. Remembering c equals 1, that will clearly have that vector divided by length. That will clearly have that vector divided by length, vector divided by length, explicitly vector divided by length. That's weird.

But remember, the basis objects themselves are a little weird. One of the things we saw was that e_ϕ has the dimensions of length. In order for the vector to itself be consistent, v_ϕ must have the dimensions of v divided by length. So in fact, when I just take its ϕ derivative, I get something that looks exactly like it should if it is to be a divergence.

Let's move on and think about how I take a covariant derivative of other kinds of tensorial objects. This is all you need to know if you are worried about taking derivatives of vectors. But we're going to work with a lot of different kinds of tensor objects.

One of the most important lectures we're going to do in about a month actually involves looking at a bunch of covariant derivatives of some four-indexed objects, so it gets messy. Let's walk our way there. So suppose I want to take the derivative of a scalar. Scalars have no basis object attached to them.

There's no basis object. When I take the derivative, I don't have to worry about anything wiggling around. No Christoffel symbols come in. If I want to take the covariant derivative of some field ϕ , it is nothing more than the partial derivative of that field ϕ -- boom. Happy days.

How about a one form? The long way to do this would be to essentially say, well, the way I started this was by looking at how my basis vectors varied as I took their derivatives. Let's do the same thing for the basis one forms, assemble my table, do a lot of math, blah, blah, blah.

Knock yourselves out if that's what you want to do. There's a shortcut. Let's use the fact that when I contract a one form on a vector, I get a scalar.

So let's say I am looking at the beta covariant derivative of p^α on a α . That's a scalar. So this is just the partial derivative.

And a partial derivative of the product of something I can expand out really easily. So using the fact that this just becomes the partial, I can write this as $a^\alpha \partial_\beta p^\alpha$ plus $p^\alpha \partial_\beta a^\alpha$. So now what?

Well, let's rewrite this using the covariant derivative. Pardon me a second while I get caught up in my notes. Here we are. I can write this as the covariant derivative minus the correction that comes from that Christoffel symbol. Pardon me just a second. There's a couple lines here I want to write out very neatly.

So when I put this in-- oops typo. That last typo is important, because I'm now going to do the relabeling trick. So what I'm going to do is take advantage of the fact that

in this last term, alpha and mu are both dummy indices. So on this last term that I have written down here, I'm going to swap out alpha and mu.

When I do that, notice that the first term and the last term will both be proportional to the component a_α . Now, let's require that the covariant derivative when it acts on two things that are multiplied together, it's a derivative, so it should do what derivatives ordinarily do. So what we're going to do is require that when I take this covariant derivative, I should be able to write the result like so.

It's a healthy thing that any derivative should do. So comparing, I look at that, and go, oh, I've got the covariant derivative of my one form there. Just compare forms.

Very, very similar, but notice the minus sign. There's a minus sign that's been introduced there, and that minus sign guarantees, if you actually expand out that combination of covariant derivatives I have on the previous line, there's a nice cancellation so that the scalar that I get when I contract p on a , in fact, doesn't have anything special going on when I do the covariant derivative.

So I'm going to generalize this further, but let me just make a quick comment here. I began this little calculation by saying, given how we started our calculation of the covariant derivative of a vector, we could have begun by just taking lots of derivatives of the basis one forms, and assembling all these various tables, and things like that. If you had done this, it's simple to find, based on an analysis like this, that if you take a partial derivative of a one form, that you get sort of a linear combination of one forms back.

Looks just like what you got when you took a partial derivative of the basis vector, but with a minus sign. And what that minus sign does is it enforces, if you go back to a lecture from ages ago, when I first introduced basis one forms, it enforces the idea that when I combine basis one forms with basis vectors, I get an identity object out of this, which is itself a constant. If you are the kind of person who likes that sort of mathematical rigor, some textbooks will start with this, and then derive other things from that-- sort of six of one, half a dozen of the other.

So we could go on at this point. And I could say, how do I do this with a tensor that has two indices in the upstairs position? How do I do this with a tensor that has two indices in the downstairs position? How do I do it for a tensor that's got 17 indices in

the upstairs position and 38 in the downstairs position? The answer is easily deduced from doing these kinds of rules, so I'm just going to write down a couple of examples and state what it turns out to be.

So basically, imagine I want to take the covariant derivative-- let's do the stress energy tensor-- covariant derivative of $T_{\mu\nu}$. So remember, the way that the Christoffel got into there is that when I looked at the derivative of a vector, I was looking at derivatives of basis objects. Well, now I'm going to look at derivatives of two different basis objects.

So I'm going to wind up with two Christoffel symbols. You can kind of think of it as coming along and correcting each of these indices. I can do this with the indices in the downstairs position. Guess what? Comes along and corrects all them with minus signs.

Just for completeness, let me just write down the general rule. If I am looking at the covariant derivative of a tensor with a gazillion upstairs indices and a gazillion downstairs indices, you get one term that's just a partial derivative of that guy, and you get a Christoffel coupling for every one of these. Plus sign for all the upstairs, minus sign for all the downstairs. That was a little tedious.

You basically just, when I give you a tensor like that, you just kind of have to go through. And it becomes sort of almost monkey work. You just have to rotely go through and correct every one of the indices using an algorithm that kind of looks like this. Oh, jeez, there's absolutely a minus sign on the second one. Thank you. I appreciate that.

So the way that we have done things so far, and I kind of emphasized, it sort of smells like the way to do this is you pick your new coordinate representation, you throw together all of your various basis objects, and then you just start going whee, let's start taking derivatives and see how all these things vary with respect to each other, assemble my table of the gammas, and then do my covariant derivative.

If that were, in fact, the way we did it, I would not have chosen my research career to focus on this field. That would suck. Certainly prior to Odin providing us with Mathematica, it would have been absolutely undoable. Even with it, though, it would be incredibly tedious.

So there is a better way to do this, and it comes via the metric. Before I derive what the algorithm actually is, I want to introduce an extremely important property of tensor relationships that we are going to come back to and use quite a bit in this course. So this is something that we have actually kind of alluded to repeatedly, but I want to make it a little more formal and just clearly state it. So this relationship that I'm going to use is some kind of a tensor equation, a tensorial equation that holds in one representation must hold in all representations.

Come back to the intuition when I first began describing physics in terms of geometric objects in spacetime. One of the key points I tried to really emphasize the difference of is that I can have different-- let's say my arm is a particular vector in spacetime. Someone running through the room at three-quarters the speed of light will use different representations to describe my arm. They will see length contractions. They will see things sort of spanning different things.

But the geometric object, the thing which goes between two events in spacetime, that does not change, even though the representation of those events might. This remains true not just for Lorentz transformations, but for all classes of transformations that we might care to use in our analysis. Changing the representation cannot change the equation. Written that way, it sounds like, well, duh, but as we'll see, it's got important consequences.

So as a warm-up exercise of how we might want to use this, let's think about the double gradient of a scalar. So let's define-- let's just say that this is the object that I want to compute. Let's first do this in a Cartesian representation.

In a Cartesian representation, I just take two partial derivatives. I've got a couple basis one forms for this. So I've got something like this. The thing which I want to emphasize is that as written, in Cartesian coordinates, $d\alpha d\beta$ of ϕ -- those are the components of a tensor in this representation.

And the key thing is that they are obviously symmetric on exchange of the indices α and β . If I'm just taking partial derivatives, doesn't matter what order I take them in. That's got to be symmetric.

Let's now look at the double gradient of a scalar in a more general representation.

So in a general representation, I'm going to require these two derivatives to be covariant derivatives. Now, we know one of them can be very trivially replaced with a partial, but the other one cannot. Hold that thought for just a second.

If this thing is symmetric in the Cartesian representation, I claim it must also be true in a general representation. In other words, exchanging the order of covariant derivatives when they act on a scalar should give me the same thing back. Let's see what this implies.

So if I require the following to be true, that's saying-- oops. So let's expand this out one more level. So now, I'm correcting that downstairs index and over here.

So the terms involving nothing but partials, they obviously cancel. I have a common factor of $d\mu$ of ϕ . So let's move one of these over to the other side.

What we've learned is that this requirement, that this combination of derivatives be symmetric, tells me something about the symmetry of the Christoffel symbols itself. If you go back to that little table that I wrote down for plane polar coordinates, that was one where I just calculated only three non-trivial components, but there was a symmetry in there. And if you go and you check it, you will see it's consistent with what I just found right here.

Pardon me for just a second. I want to organize a few of my notes. These have gotten all out of order. Here it is.

So let me just use this as an opportunity to introduce a bit of notation. Whenever I give you a tensor that's got two indices, if I write parentheses around those indices, this is going to mean that I do what is called "symmetrization." We're going to use this from time to time.

If I write square braces, this is what we call "anti-symmetrization." And so what we just learned is that $\gamma_{\mu\alpha\beta}$ is equal to $\gamma_{\mu\alpha\beta}$ with symmetrization on those last two indices. We have likewise learned that if I contract this against some object, if these were anti-symmetric, I must get a 0 out of it. So that's a brief aside, but these are important things, and I want to make sure you have a chance to see them.

So trying to make a decision here about where we're going to want to carry things forward. We're approaching the end of one set of notes. There's still one more thing I want to do.

So I set this whole thing up by saying that I wanted to give you guys an algorithm for how to generate the Christoffel symbols. The way I'm going to do this is by examining the gradient of the metric. So suppose I want to compute the following tensor quantity-- let's say is g the metric tensor, written here in the fairly abstract notation. And this is my full-on tensor gradient of this thing.

So if you want to write this out in its full glory, I might write this as something like this. But if you stop and think about this for just a second, let's go back to this principle. An equation that is tensorial in one representation must be tensorial in all. Suppose I choose the Cartesian representation of this thing.

Well, then here's what it looks like there. But this is a constant. So if I do this in Cartesian coordinates, it has to be 0. The only way that I can make this sort of comport with this principle that an equation that is tensorial in one representation holds in all representations-- this leads me to say, I need to require that the covariant derivative of the metric be equal to 0.

We're going to use this. And I think this will be the last detailed calculation I do in today's lecture. We're going to use this to find a way to get the Christoffel symbol from partial derivatives of the metric.

There's a lot of terms here and there's a lot of little indices. So I'm going to do my best to make my handwriting neat. I'm going to write down a relationship that I call "Roman numeral I." The covariant derivative in the γ direction, $G_{\alpha\beta}$ -- you know what, let me put this down a little bit lower, so I can get these two terms on the same line.

So I get this thing that involves two Christoffel symbols correcting those two indices. This is going to equal 0. I don't really seem to have gotten very far. This is true, but I now have two bloody Christoffel symbols that I've somehow managed to work into this.

What I'm trying to do is find a way to get one, and equate it to things involving

derivatives of the metric. So this is sort of a ruh-roh kind of moment. But there's nothing special about this order of the indices.

So with the audacity that only comes from knowing the answer in advance, what I'm going to do is permute the indices. Then go, oh, let's permute the indices once more. So I'll give you guys a moment to catch up with me.

Don't forget, these notes will be scanned and added to the web page. So if you don't want to follow along writing down every little detail, I understand, although personally, I find that these things gel a little bit better when you actually write them out yourself.

So those are three ways that I can assert that the metric has no covariant derivative. They all are basically expressing that same physical fact. I'm just permuting the indices.

Now there's no better way to describe this than you sort of just stare at this for a few moments, and then go, gee, I wonder what would happen if-- so stare at this for a little while. And then construct-- you know I have three things that are equal to 0.

So I can add them together, I can subtract one from the other. I can add two and subtract one, whatever. They should all give me 0.

And the particular combination I want to look at is what I get when I take relationship one and I subtract from it two and three. So I'm going to get one term that are just these three combinations of derivatives, gamma. And I get something that looks like-- let me write this out and then pause and make a comment.

So I sort of made some lame jokes a few moments ago that essentially, the only reason I was able to get this was by knowing the answer in the back of the book, essentially. And to be perfectly blunt, for me personally, that's probably true. When I first wrote this down, I probably did need to follow an algorithm.

But if I was doing this *ab initio*, if I was sitting down to first do this, what's really going on here is the reason I wrote out all these different combinations of things is that I was trying to gather terms together in such a way that I could take advantage of that symmetry. A few moments ago, we showed that the Christoffel symbols are symmetric on the lower two indices.

And so by putting out all these different combinations of things, I was then able to combine them in such a way that certain terms-- look at this and go, ah, symmetry on alpha and gamma means this whole term dies. Symmetry on beta and gamma means this whole term dies. Symmetry on alpha and beta means these two guys combine, and I get a factor of 2.

So let's clean up our algebra. Move a bunch of our terms to the other side equation, since it's a blah, blah, blah equals 0. And what we get when we do this is $g_{\mu\nu}$ downstairs gamma is equal to $1/2$.

What we're going to do now is we will define everything on the right-hand side-- I've kind of emphasized earlier that the Christoffels are not themselves tensors, but we're going to imagine that we can nonetheless-- we're not going to imagine, we're just going to define-- we're going to say that we're allowed to raise and lower their indices using the metric, in the same way you guys been doing with vectors and one forms and other kinds of tensors. So let's call everything on the right-hand side here gamma with all the indices in the downstairs position, $\gamma_{\mu\nu}$. And then this is simply what I get when I click all of these things together like so.

If you go and you look up the formulas for this in various textbooks that give these different kinds of formulas, you will typically see it written as $1/2 g^{\mu\nu}$ upstairs indices, and then all this stuff in parentheses after that. When you look things up, this is the typical formula that is given in these books. This is where it comes from.

So I need to check one thing because it appears my notes are a little bit out of order here. But nonetheless, this is a good point, since we've just finished a pretty long calculation, this is a good point to introduce an important physical point. We're going to come back to this. We're going to start this on Thursday.

But I want to begin making some physical points that are going to take us from special relativity to general relativity. So despite the fact that I've introduced this new mathematical framework, everything that I have done so far is in the context of special relativity. I'm going to make a more precise definition of special relativity right now.

So special relativity-- we are going to think of this moving forward as the theory which allows us to cover the entire spacetime manifold using inertial reference frames. So we use inertial reference frames or essentially, Lorenz reference frames, and saying that Lorenz coordinates are good everywhere. We know we can go between different Lorenz reference frames using Lorenz transformations.

But the key thing is that if special relativity were correct, the entire universe would be accurately described by any inertial reference frame you care to write down. And I will probably only be able to do about half of this right now. We'll pick it up next time, if I cannot finish this.

The key thing which I want to emphasize is, gravity breaks this. As soon as you put gravity into your theory of relativity, you cannot have-- so we will call this a global inertial frame, an inertial frame that is good everywhere, so "global" in the mathematical sense, not "global" in the geographic sense-- not just earth, the whole universe, essentially. As soon as we put in gravity, we no longer have global reference frames and global inertial reference frames.

That word "inertial" is important. But we are going to be allowed to have local inertial frames. I have not precisely defined the difference what "local" means in this case, and I won't for a few lectures.

But to give you a preview as to what that means, it's essentially going to say that we can define an inertial coordinate system that is good over a particular region of spacetime. And we're going to have to discuss and come up with ways of understanding what the boundaries of that region are, and how to make this precise. So the statement that gravity breaks the existence of global Lorenz frames, like I said, it's a two-part thing.

I'm going to use a very handwavy argument which can be made quite rigorous later, but I want to keep it to this handwaving level, because first of all, it actually was first done by a very high-level mathematical physicist named Alfred Schild, who worked in the early theory of relativity. It's sort of like he was so mathematical, if it was good enough for him, that's good enough for me. And I think even though it's a little handwavy, and kind of goofy in at least one place, it gives a good physical sense as to why it is gravity begins to mess things up.

So part one is the fact that there exists a gravitational redshift. So here's where I'm going to be particularly silly, but I will back up my silliness by the fact that everything silly that I say here has actually been experimentally verified, or at least the key physical output of this. So imagine you are on top of a tower and you drop a rock of rest mass m off the top of this tower.

So here you are. Here's your rock. The rock falls.

There's a wonderful device down here which I label with a p . It's called a photonulator. And what the photonulator does, it converts the rock into a single photon, and it does so conserving energy.

So when this rock falls, the instant before it goes into your photonulator, just use Newtonian physics plus the notion of rest energy. So it's got an energy of m -- $m c^2$ squared, if you prefer, its rest energy-- plus what it acquired after falling-- pardon me, I forgot to give you a distance here-- after falling a distance h . So that means that the photon that I shoot up from this thing-- let me put a few things on this board.

So the instant that I create this photon, this thing goes out, and it's going to have a frequency ω bottom, which is simply related to that energy. This photon immediately is shot back up to the top, where clever you, you happen to have in your hands a rerockulator. The rerockulator, as the name obviously implies, converts the photon back into a rock.

Now, suppose it does so-- both the photonulator and the rerockulator are fine MIT engineering. There are no losses anywhere in this thing. So there's no friction. There's no extra heat generated. It does it conserving energy, 100%.

What is the energy at the top? Well, you might naively say, ah, it's just going to go up to the top. It's going to have that same energy. It might just have it in the form of a photon and ω b.

There will be some frequency at the top. And your initial guess might be it's going to be the same as the frequency at the bottom. But if you do that, you're going to suddenly discover that your rock has more energy than it started out with, and you can redirect it back down, send it back up. Next thing you know, you've got yourself

a perpetual motion machine.

So all you need to do is get your photonulater and your rerockulater going, and you've got yourself a perpetual motion machine here. I will grant that's probably not the weakest part of this argument. Suppose you had this.

I mean, you look at this. If technology allowed you to make these goofy devices, you would instantly look at this and say, look, if I am not to have-- let's just say I live in a universe where I'm fine with photonulaters. I'm fine with rerockulaters, but damn it, energy has to be conserved. I am not fine with perpetual motion machines.

If that's the case, we always fight perpetual motion. We must have that the energy at the top is equal to the energy this guy started with. When it sort of gets back into-- imagine that your rerockulater is shaped like a baseball catcher's mitt. You want that thing to just land gently in your mitt, and just be a perfectly gentle, little landing there.

And when you put all this together, going through this, taking advantage of the fact that if you work in units where you've put your c's back in, there will be a factor of g over c squared appearing in here, what you find is that the frequency at the top is less than the frequency at the bottom. In other words, the light has gotten a little bit redder. now I fully confess, I did this via the silliest argument possible.

But I want to emphasize that this is one of the most precisely verified predictions of gravity and relativity theory. This was first done, actually, up the street at Harvard, by what's called the Pound-Rebka experiment. And the basic principles of what is going on with this right now-- I just took this out to make sure my alarm is not about to go off, but I want to emphasize it's actually built into the workings of the global positioning system.

Because this fact that light signals may travel out of a gravitational potential, they get redshifted, needs to be taken into account in order to do the precise metrology that GPS allows. Now, this is part one, this idea that light gets redder as it climbs out of a gravitational field. Part two, which I will do on Thursday, is to show that if there is a global inertial frame, there is no way for light to get redder as it climbs out of a gravitational potential.

You cannot have both gravity and a global inertial reference frame. That's where I will pick it up on Thursday. So we'll do that.

And we will then begin talking about how we can try to put the principles of relativity and gravity together. And in some sense, this is when our study of general relativity will begin in earnest. All right, so let us stop there.