

Intuition and Primer for General Relativity and Differential Geometry

Ben Pennell

V1.0.4 August 28th, 2025

Abstract

My aim with this document is to bridge the gap that I often see (and experienced myself) when physicists are taught cosmology. In a perfect world, courses building to and on differential geometry would be taken before courses on General Relativity and Cosmology. For physics undergraduates, this is often not possible and in all honesty is not required. However, this makes learning the material more challenging and courses are often not designed with this in mind. In this document I will start from early undergraduate physics and mathematics taken in standard physics degrees. In other words I will only require the following knowledge: multivariable calculus, linear algebra, ordinary differential equations, and some classical mechanics (if you know what a Lagrangian is, you're set). From that, I will motivate many critical elements for the machinery and results normally taught in courses on General Relativity and Cosmology.

Contents

1	Additional Reading	1
2	Pythagoras to Einstein	2
2.1	The line element	2
2.2	Special Relativity	3
3	What the hell is a tensor?	5
3.1	Manifolds	5
3.2	Vectors: Covariance and Contravariance	6
3.3	Vector transformations	7
3.4	Building to the Covariant Derivative	7
3.5	Tensors	9
3.6	The Metric Tensor	9
4	Fields	10
4.1	Note on Counting Degrees of Freedom	10
4.2	What the hell is an action?	11
4.3	Noether's Theorem	13
4.4	Intuition From Electromagnetism (HEL 6)	15
5	Gravitational Waves (HEL 17&18)	16
5.1	Perturbations	17
5.2	Obtaining the field equations	17
5.3	Stretchy-Squeezy	18

1 Additional Reading

Here are the books I've found to be the most useful when I was learning this material. Each has different things to offer, and this document is essentially a compilation of sections of each of these that worked best for my understanding. Full versions can be found online for all of these works.

1. Matthias Bartelmann: General Relativity
2. Hobson MP, Efstathiou GP, Lasenby AN: General Relativity: An Introduction for Physicists
 - Great introduction to the mathematics for physicists. Is largely what these notes are based on. Sections significantly inspired will be denoted by HEL and chapters will be listed

3. David Tong: Lectures on General Relativity (online) <https://www.damtp.cam.ac.uk/user/tong/gr.html>
4. Daniel Baumann: Cosmology
5. Landau, Lifshitz: Course of Theoretical Physics volume 2: The Classical Theory of Fields
 - Great introduction to the principles of special relativity, and provides the best understanding of field theories and the action. As a warning, this book is challenging to read and requires serious time commitment

2 Pythagoras to Einstein

2.1 The line element

We will start ourselves off as simple as we possibly can. If we're working in 3-dimensional euclidean space, you can calculate the distance (for the sake of convention, written as s) between two points with

$$s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 \quad (1)$$

This is *Pythagoras' theorem*, I sure hope you're familiar with it.

Now, let's consider two points that are infinitesimally separated. We just change our finite difference Δ to an infinitesimal d

$$ds^2 = dx^2 + dy^2 + dz^2 \quad (2)$$

Still nice and simple. Now, let's bring up the single most important fact that you must keep in mind for the rest of your GR career

Physics is not coordinate-dependent

(3)

Physics works as it does, regardless of what you choose to call the center of your system, or how big you define a meter or a second to be. My GR professor, David Curtin's words have been drilled into my mind: *if you find that your physics is coordinate-dependent, it is not interesting, it is wrong*. This is a good thing to keep in mind, and is actually a good sanity check to make sure you've done calculations correctly.

Let's return to our distance between points.

We could have also worked in spherical coordinates, and physics should remain the same. You can work this out for yourself if you'd like, but the distance between two infinitesimally separated points would be

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2(\theta) d\phi^2 \quad (4)$$

Critically, ds^2 is a physical object. Whether defined by spherical, euclidean, or any other coordinate system, it is the same object in space, we just give it different representations.

There must be a way that we can write ds^2 that doesn't depend on coordinates, then. This would be nice to have, since we wouldn't have to worry about our choice of coordinates. We can work this out by finding a way to generate both Equations 2 and 4 from a single functional form to find this general coordinate-free representation. To start, we could have also written ds^2 in the euclidean coordinates as

$$ds^2 = \sum_i dx_i^2 \quad (5)$$

Where we have three different components $x_1 = x, x_2 = y, x_3 = z$. Likewise, we could write ds^2 with the spherical coordinates as

$$ds^2 = \sum_i c_i dx_i^2 \quad (6)$$

Where we have a set of constants $c_1 = 1, c_2 = r^2, c_3 = r^2 \sin^2(\theta)$ and coordinates $x_1 = r, x_2 = \theta, x_3 = \phi$. In fact, this form also works for euclidean space, where each of $c_i = 1$.

We can simplify our notation a bit. Henceforth, if you ever see an index (like i in the previous equation) that appears twice in one term, it's implied to be summed over. In general, one writes one of the indices as a superscript and one as a subscript. You can spot the summation by matching a superscript and subscript index. Our equation would take the form

$$ds^2 = c^i dx_i^2 \quad (7)$$

Which means the same thing. We soon will see these upper and lower indices generalize to represent *covariant* and *contravariant* components of vectors. But for now, we'll carry on

Now, if I had some space that was more complicated than euclidean space, I could have some other weird features in ds^2 . For example, a rotating wormhole could have it look like

$$ds^2 = A dt^2 + B dr^2 + C d\theta^2 + D d\phi^2 + E dt d\phi \quad (8)$$

Where this comes from isn't too relevant, but for now, notice that we have a term with $dt d\phi$. So, it is possible that we won't just have all of our coordinates showing up quadratically. If we want to extend our definition of ds^2 to incorporate this, we need to get slightly more complicated.

We will have to promote our 1-dimensional vector c^i to a 2-dimensional matrix. You can think of this as the coefficient for every possible mix of coordinates. For the rotating wormhole, it would look like

$$\begin{bmatrix} A & 0 & 0 & E \\ 0 & B & 0 & 0 \\ 0 & 0 & C & 0 \\ E & 0 & 0 & D \end{bmatrix} \quad (9)$$

While for spherical coordinates it would look like

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2(\theta) \end{bmatrix} \quad (10)$$

In Cartesian, it is simply the identity matrix. For now, you can also note that this matrix has to be symmetric, since swapping the order of multiplication makes no difference.

The convention is to call this matrix $g_{\mu\nu}$, where the labels μ and ν tell you which row and column to look at, respectively.

Our equation for ds^2 now takes the form

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (11)$$

Noting that we sum over every possible index μ and ν . This could just as easily be written

$$ds^2 = g^{\mu\nu} dx_\mu dx_\nu \quad (12)$$

And now this represents our object ds^2 for *any* coordinate system you want!

Now, we can define some differential geometry lingo. ds^2 is referred to as the *line element*, and $g_{\mu\nu}$ is the *metric*. Additionally, $g_{\mu\nu}$ is our first example of a *tensor*. We'll explore what a tensor even is in the next chapter.

In general, the metric will define a volume element in space. In fact, it is the square root of the determinant of this matrix that defines it. This can be made clear by taking the determinant of the spherical coordinate matrix drawn above. You get $R^4 \sin^2(\theta)$, the square root of which gives $R^2 \sin(\theta)$ which is the volume element that you may recall from multivariable calculus.

2.2 Special Relativity

Now, let's get into the specifics of special relativity. First, we take from experiments that the *principle of relativity* is valid. That is, the laws of physics are the same in all inertial frames of reference. When we do physics in our classrooms, like for example classical mechanics with interacting particles, we assume the particles interact instantaneously. We take from observations that instantaneous interactions do not exist. Since there are no instantaneous interactions, there is some maximum *signal velocity* at which interactions propagate. This is because some amount of time has to elapse between an event and another body witnessing it, you can divide the distance by that time and work out the velocity that the signal travelled at. Notably, this must simply be a fundamental feature of the vacuum.

As a result of the principle of relativity, and the fact that this is a property of the vacuum and not any particular object, this signal velocity should be the same in all inertial frames of reference and is a universal constant. We call this velocity *the speed of light*. One can actually see how this velocity is a natural property of the vacuum as a result of electromagnetism.¹

Consider the maxwell equations in free space, you can find that they generate propagating wave solutions. I'll leave it as an exercise, but by decoupling the E and B equations by cleverly applying the curl and divergence operators, that for the electric field look like

¹although the speed of light is *the* fundamental constant. Permittivity and Permeability exists as a result of *it*. I demonstrate this example in this way because it helped me connect disparate aspects of physics together in my mind.

$$\nabla^2 E = \mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2} \quad (13)$$

We will also see how to derive this functional form in Section 4.4. For now, just know that this simply describes a wave that travels at a speed related to the coefficient $v = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$, which of course is the speed of light!

Recall that ϵ_0 and μ_0 are just constants (properties of the vacuum) that you used in your electrostatics formulas, and their values were determined from observations, and the speed of light is a constant based on them.

The next insight to build into special relativity is to consider time as a coordinate. After all, we're all moving forwards in time at a (usually) constant rate. Of course, it may seem odd to think of time as a coordinate alongside space, but by multiplying time by the speed of light, you get a distance and so the units make sense. One would write the line element for *flat spacetime* or *Minkowski spacetime* like

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \quad (14)$$

And, in general, one might also swap all the signs on the right hand side of this equation if they so chose. Choosing where to put the negative signs is just a matter of notation and is called the *signature* of the space. For this document, I will take the $(-+++)$ convention because it's totally better.

What this line element describes is the difference between your spacial distance and the signal distance travelled in the same amount of time. As a result, light travels on *null* paths, where $ds^2 = 0$. So, the magnitude of the line element of Minkowski space just describes your deviation from a light path. One would call a positive line element *spacelike* and a negative line element *timelike*. For normal situations, things are spacelike.

Of course, since one cannot exceed the speed of light, the addition of velocities takes a more complicated form to avoid the fact that standard summing of velocities could cause you to end up faster than light.

Consider euclidean space with just two coordinates: time t and one spacial coordinate x . We take two frames of reference: one which is *boosted* along the x -direction by a velocity v . We denote the boosted frame with primes. Furthermore, the two frames of reference both start at $t = t' = 0$.

In general, the new coordinates are arbitrary functions of the old coordinates

$$t' = At + Bx \quad (15)$$

$$x' = Dt + Ex \quad (16)$$

The game here now is to constrain the functional forms and values of A, B, D, E .

Since the primed coordinates are just a boost of the original coordinates, we have constructed that at $x' = 0$, $x = vt$. So, we can write the second equation from above as

$$0 = Dt + Evt \quad (17)$$

Which yields that $D = -Ev$. Similarly, at $x = 0$, $x' = -vt'$, so, we can rewrite the first equation as

$$t' = At \quad (18)$$

Which can then be plugged into the second equation,

$$-vt' = -Evt \quad (19)$$

$$\implies -vAt = -Evt \quad (20)$$

Meaning that $A = E$. We can rewrite the primed coordinates as

$$t' = At + Bx \quad (21)$$

$$x' = A(x - vt) \quad (22)$$

Now, one can imagine writing the line element in both coordinate systems

$$ds^2 = -c^2 t'^2 + x'^2 = -c^2 t^2 + x^2 \quad (23)$$

We can plug in the relationships from above,

$$-c^2t^2 + x^2 = -c^2(At + Bx)^2 + (Ax - Avt)^2 \quad (24)$$

$$-c^2t^2 + x^2 = -c^2A^2t^2 - c^2B^2x^2 - 2c^2AtBx + A^2x^2 + A^2v^2t^2 - 2A^2xvt \quad (25)$$

So, from dimensional analysis we can collect the t^2 terms

$$-c^2t^2 = -c^2A^2t^2 + A^2v^2t^2 \quad (26)$$

$$\downarrow \quad (27)$$

$$A = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \equiv \gamma \quad (28)$$

Where we call γ the *lorentz factor*. Also from dimensional analysis, we can collect the tx terms

$$2c^2AtBx = -2A^2xvt \quad (29)$$

$$B = -\frac{v}{c^2}A \quad (30)$$

So, finally, we obtain

$$t' = \gamma\left(t - \frac{v}{c^2}x\right) \quad (31)$$

$$x' = \gamma(x - vt) \quad (32)$$

Which is called a *lorentz boost* in the x -direction. this factor of γ prevents speeds from crossing the speed of light. This comes into play with two effects which can be read off from the previous equations for the boosted coordinates:

- Length Contraction: a proper length l along the boost gets reduced to $\frac{l}{\gamma}$
- Time Dilation: the time observed by a particle travelling with the boost is increased to γt

We can now amend the principle of relativity

The laws of physics are the same in all intertial frames of reference, and those laws are that of special relativity (33)

3 What the hell is a tensor?

Oh, you'd like to know what a tensor is? Here's the most precise possible answer:

- Tensors are objects that transform as tensors

Does that help?

I imagine not. I hope that by the end of this section, you will not only understand this circular sounding argument, but you will agree that it's actually the best and most intuitive way to talk about tensors.

3.1 Manifolds

Let's start things off by introducing an important (and not too complicated) piece of differential geometry, the concept of a *manifold*.

A mathematician may define a manifold like

A manifold is a parameterisable surface whose local geometry is euclidean

Which is a bit of word soup, so let me translate that.

Think of the surface of the earth. It is a surface on which I can create coordinate systems (like, longitude and latitude) to define my position uniquely on it, this is it being "parameterisable".² Further, while I'm a

²Of course, you can't actually define it with only one coordinate system. To understand why, think: what is your longitude when you're at the north pole?

small human standing on the surface of the earth, it looks like it's flat (but rest assured, it's not). When I am calculating how far I throw a baseball, I can get a pretty good approximation by ignoring the earth's curvature. This is it being "locally euclidean".

We've encountered manifolds in our physics careers many times before. In classical or statistical mechanics, you may have encountered phase space (q_i, p_i) . For a single particle, this is a six-dimensional manifold. For a collection of N particles in statistical mechanics, this is a $6N$ -dimensional manifold.

We can further define a *differentiable manifold* as

- continuous: there exist infinitesimally separated points in the neighbourhood of any given point p on the manifold
- differentiable: a differentiable scalar field can be established everywhere

Which is some jargon, but just understand that this is just the same definition of a continuous function from first year calculus, just with some more grown-up verbiage to be more general. This also never really comes up in GR, all of our manifolds will be well behaved.

Finally, for interests' sake, one could write line elements that are more complicated than what we've seen, for manifolds that are more complicated than we've seen, but this isn't necessary for physics applications. We stick to the form described before, which is called a *pseudo-Riemannian manifold*. This just requires (essentially) that all terms in the line element are quadratic.

$$ds^2 = g^{\mu\nu} dx_\mu dx_\nu \quad (34)$$

3.2 Vectors: Covariance and Contravariance

Vectors are objects which can be labelled with one index. You can describe them with several components that each scale the basis vectors e_i of your coordinate system.

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_1 e_1 + v_2 e_2 + v_3 e_3 = v^i e_i \quad (35)$$

So, if you change your coordinate system, you would change e_i , and so you would get different values for v_i . But, like with the line element, the object \vec{v} exists regardless of how you define your coordinates.

Now, we formalize the nature of these upstairs and downstairs indices. Standard vectors are written as $v^i e_i$. But, we can define so-called *dual basis vectors* e^i with

$$e^i e_j = \delta_j^i \quad (36)$$

Which are a complementary set of coordinates to the original. Writing our vector \vec{v} in this dual basis is done as

$$\vec{v} = v_i e^i \quad (37)$$

Formally, we refer to these descriptions as

- v^i is the contravariant component
- v_i is the covariant component

A way to internalize what contravariant and covariant mean is to consider a transformation. If we have a vector that exists in space, and we change the length of one of the basis vectors, the value of the corresponding component of the vector will change. Covariant components change with the transformation, while contravariant components change inversely to the transformation. This makes their naming scheme intuitive, "co" means with so covariant is changing-with, while "contra" means against so contravariant means changing-against.

As an example, if I have a vector in euclidean space, and then I double the length of my x basis vector, the covariant representation of the x coordinate of my vector would double, while the contravariant representation of the x coordinate would half. For now it may seem arbitrary to use these, but you will see how they affect the math and are unfortunately quite important.

Here's another way to understand the difference. The dot product of two vectors v and w is defined as

$$v \cdot w = v^i w_i \quad (38)$$

So you can in some way think of the contravariant component as a row vector, while the covariant is a column vector. After all, to do that matrix multiplication in the linear algebra sense, you'd have to set up your vectors that way. So, back in first year when you'd arbitrarily switch a column vector to a row vector to do a dot product, you were actually jumping between contravariant and covariant representations!

3.3 Vector transformations

Let's think about how vectors transform when you change coordinates. Imagine you have two coordinate systems. A vector \vec{v} is a physical object which can be defined with either coordinate system but physically remain unchanged, so for different representations a and b ,

$$\vec{v} = v^a e_a = \tilde{v}^b \tilde{e}_b \quad (39)$$

Isolating for the new vector representations \tilde{v}^b ,

$$\tilde{v}^b = \frac{e_a}{\tilde{e}_b} v^a \quad (40)$$

This tells us how the new vector looks after the transformation. Now, let's consider the line element. This is also a physical object that can be represented by either coordinate system

$$ds = e_a dx^a = \tilde{e}_b d\tilde{x}^b \quad (41)$$

In other words,

$$\frac{e_a}{\tilde{e}_b} = \frac{d\tilde{x}^b}{dx^a} \quad (42)$$

Which we can plug into our equation for \tilde{v}^b ,

$$\tilde{v}^b = \frac{d\tilde{x}^b}{dx^a} v^a \quad (43)$$

Which tells you how vectors change with a change in coordinates.

$\frac{d\tilde{x}^b}{dx^a}$ can be written as a matrix, and this is known as the *Jacobian*. Does that name sound familiar? *I hope it does.* Just like when defining the line element before, and we had a volume element show up, we now have to define how this volume element changes. This is just like when you made coordinate transforms in multivariable calculus or elsewhere, and had to use a Jacobian matrix (and the square root of the determinant of it) to think about how the volume element changes when you computed integrals in different coordinate systems. It's the same principle here.

But notice how this is only telling us how the contravariant components transform. How do the covariant ones transform? We can work this out by starting with swapping all the indices from the contravariant relationship

$$\tilde{v}_b = \frac{e_a}{\tilde{e}^b} v^a \quad (44)$$

If we want to have the same structure in our Jacobian (that is, all upper indices in our coordinates dx^i), we have to work with lower indices in our basis vectors. Recall that the basis vectors and their duals are related by $e^i = \frac{1}{e_i}$. So,

$$\tilde{v}_b = \frac{\tilde{e}_b}{e_a} v^a \quad (45)$$

Which gives us,

$$\tilde{v}_b = \frac{dx^a}{d\tilde{x}^b} v_a \quad (46)$$

Which you can see is the inverse of the Jacobian for the contravariant case. This is the mathematical way of describing what I was talking about before, where the covariant and contravariant component scale inversely to each other.

3.4 Building to the Covariant Derivative

Euclidean space is super simple, but once we even jump to spherical coordinates there's some other weird effects. Think about setting up the surface of a sphere in spherical coordinates. As you move latitudinally, the direction of your $\vec{\theta}$ basis vector changes direction, and as you move longitudinally, your $\vec{\phi}$ basis vector changes! that's super weird, and there must be a mathematical way to describe this.

We consider how a basis vector changes as you move in a particular direction. The derivative of a basis vector with respect to a particular coordinate is, in general, dependent on every other coordinate. So, you could colloquially say that all the basis vectors are connected to one another and the functional form of this relationship is their *connection*.

$$\frac{\partial e_b}{\partial x^c} = \Gamma_{bc}^a e_a \quad (47)$$

This equation looks complicated but all I've done is define this object Γ_{bc}^a that relates every basis vector to the derivatives of every other basis vector by every coordinate. Γ_{bc}^a goes by many names. In defining it I called it *the connection*. More technically it could be called the *levi-civita connection* or the *affine connection*, which is the name I prefer to use.³ Often it is also called a *Christoffel*, but that's an unenlightening name.

Rearranging, we get the formula for the affine connection

$$\Gamma_{bc}^a = e^a \frac{\partial e_b}{\partial x^c} \quad (48)$$

This is useful, because taking the derivative of a vector is actually non-trivial. Let's say I wanted to see how my vector field \vec{v} changes as I move along a particular direction. This is the following derivative, which we will write in a compact notation

$$\frac{\partial \vec{v}}{\partial x^a} = \partial_a \vec{v} \quad (49)$$

Recall that a vector is actually the dot product between its components in a particular coordinate system and the basis vectors in that coordinate system. You'll actually have to product rule it and you'll have two terms

$$\partial_a \vec{v} = \partial_a (v^b e_b) = e_b \partial_a v^b + v^b \partial_a e_b \quad (50)$$

for the second term, we will need our affine connection! Good thing we've defined that,

$$\partial_a \vec{v} = e_b \partial_a v^b + v^b \Gamma_{ab}^c e_c \quad (51)$$

We are free to relabel summed-over indices, doing so let's you write this as

$$\partial_a \vec{v} = (\partial_a v^b + v^c \Gamma_{ac}^b) e_b \quad (52)$$

the inside of the bracket is defined as the *covariant derivative*.

$$\nabla_a v^b = \partial_a v^b + v^c \Gamma_{ac}^b \quad (53)$$

This might seem arbitrary, but this should actually make logical sense. This is just a silly looking version of the multivariable chain rule! Think about it,

$$\frac{d}{dt} f(x(t), t) = \frac{\partial}{\partial t} f + \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} \quad (54)$$

The total derivative of this function has two terms: one with the simple partial derivative, and one where you chain deeper into the function. These are analogous to the two terms in the covariant derivative. Elsewhere in physics you see this, from the convective derivative to Liouville's theorem. So, it shouldn't seem that crazy. The covariant derivative can be just thought of as a total derivative of a vector, in that sense, it's obvious we need a connection term.

What's particularly powerful about the covariant derivative is that we've removed coordinate dependence. So,

If a covariant equation is true in one coordinate system, it's true in all coordinate systems

(55)

Combining this with an abuse of symmetry allows us to do most of what we do in general relativity. So, it will be the general strategy going forward to move equations into "Covariant" mode by using covariant derivatives and other coordinate-free representations.

³For reasons beyond this document, I find this to be the most illuminating way to name this as it connects to this adjective *affine* which is used elsewhere in maths.

3.5 Tensors

Let's take stock of the objects that we've come across

- A scalar is a zero-index object s
- A vector is a one-index object, like a velocity v^i
- A tensor is a two-index object, like our metric $g_{\mu\nu}$

We can return to that circular argument from the beginning of this section and understand it a bit better. Think about our line element, represented in two coordinate systems

$$ds^2 = g_{ab}dx^a dx^b = \tilde{g}_{cd}d\tilde{x}^c d\tilde{x}^d \quad (56)$$

To figure out how the metric transforms with the new coordinate system, we want to solve for \tilde{g}_{cd} in terms of g_{ab} . We can write into this equation,

$$g_{ab} \left(\frac{dx^a}{d\tilde{x}^c} d\tilde{x}^c \right) \left(\frac{dx^b}{d\tilde{x}^d} d\tilde{x}^d \right) = \tilde{g}_{cd} d\tilde{x}^c d\tilde{x}^d \quad (57)$$

And by rearranging, we get

$$\left(g_{ab} \frac{dx^a}{d\tilde{x}^c} \frac{dx^b}{d\tilde{x}^d} \right) d\tilde{x}^c d\tilde{x}^d = \tilde{g}_{cd} d\tilde{x}^c d\tilde{x}^d \quad (58)$$

We can read off that

$$\tilde{g}_{cd} = g_{ab} \frac{dx^a}{d\tilde{x}^c} \frac{dx^b}{d\tilde{x}^d} \quad (59)$$

This is how a tensor transforms under such a coordinate change. In general, anything is a tensor if they transform using this equation, just as we discussed at the beginning of the section. You can think of it as being scaled by two jacobians, one for each index. For higher order tensors (ones with three or more indices), you would just pick up more jacobians.

3.6 The Metric Tensor

To promote ourselves from special relativity (Minkowski spacetime) to general relativity (curved spacetime), we need to get a better handle on the metric tensor.

The metric tensor was used to define a line element on a manifold, like,

$$ds^2 = g^{\mu\nu} dx_\mu dx_\nu \quad (60)$$

and notably the metric in a way defines the surface, so one might expect that it is a function of the basis vectors which span that surface. Let's think more about what basis vectors actually are.

Imagine two infinitesimally separated points, and a vector \vec{s} pointing from one to the other. We have a set of coordinates x^i . We would define the basis vector e_i corresponding to each coordinate to be the vector tangent to the curve \vec{s} , along x^i , that define an infinitesimal separation. In mathematical terms, this is written as

$$e_i = \lim_{\delta x^i \rightarrow 0} \frac{\delta \vec{s}}{\delta x^i} \quad (61)$$

One could immediately see that this can be rewritten as (which an interchange of displacement δ and infinitesimal d)

$$ds = e_i dx^i \quad (62)$$

which, squared, is just the line element

$$ds^2 = e_i e_j dx^i dx^j \quad (63)$$

From which we can read off that

$$g_{\mu\nu} = e_\mu \cdot e_\nu \quad (64)$$

So, the metric is directly a function of our basis vectors.

The metric has another interesting property that has until now been glossed over. It can be used to raise and lower indices in other objects. Think about it this way: for a vector \vec{v} ,

$$\vec{v} = v^a e_a \quad (65)$$

Which, with the identity that $e^a e_b = \delta_b^a$, can be rearranged to

$$\vec{v} e^a = v^a \quad (66)$$

now, consider

$$g_{ab} v^a \quad (67)$$

What would this become? Well, we can take our above equation and substitute it in

$$g_{ab} v^a = g_{ab} \vec{v} e^a \quad (68)$$

Now, using the basis vector representation of the metric,

$$g_{ab} \vec{v} e^a = e_a e_b \vec{v} e^a = e_b \vec{v} = v_b \quad (69)$$

So,

$$g_{ab} v^a = v_b \quad (70)$$

So, the metric just switches the index from upstairs to downstairs! This applies just as equally to tensors, and is invaluable in doing calculations in General Relativity.

Armed with this knowledge, one can rewrite the affine connection that was established in the previous section, and write it in terms of the metric. This is a long calculation and isn't particularly illuminating, but the result is

$$\Gamma_{bc}^a = \frac{1}{2} g^{ad} (\partial_b g_{dc} + \partial_c g_{bd} - \partial_d g_{bc}) \quad (71)$$

So, it's a bunch of derivatives of the metric! Recall its motivation as part of the definition of the covariant derivative. In essence, it's describing the intrinsic *slope* of the manifold. The way we use it is indeed quite similar to how one uses the first derivative while analysing functions.

4 Fields

Field theory is hard. As someone from the middle of the Canadian prairies, one might think that I'd understand fields well, but unfortunately I found this material very challenging to learn. The notation is what makes it especially difficult, so I will build up different parts of the notation and pieces of intuition in separate sections. Hopefully it helps our understanding when taken as a whole.

4.1 Note on Counting Degrees of Freedom

It is sometimes a challenge to understand how many degrees of freedom we're working with once we jump into field theories, and the notation can very much obscure it. To build some better intuition, let's start all the way from the bottom to understand how to spot them. Let's start with the dead simple. In the following equation,

$$A + B + C = D \quad (72)$$

There are three degrees of freedom. If I then add another equation to the system,

$$2A - 3B + C = 0 \quad (73)$$

We end up with two degrees of freedom, since one has been eliminated, because the second equation represents a new constraint. Formally, this is because this is linearly independent of our previous equation, and so adds an additional constraint on the values of A , B , and C . So, one has to be careful because you could have what looks like four constraining equations, but they may only represent three (or fewer) degrees of freedom if they are not independent of each other.

If I then add two more independent equations to our system, each of these variables would be uniquely defined (in other words, no more degrees of freedom). We understand how this works, each equation eliminates a degree of freedom, and a number of equations equal to the number of variables uniquely identifies them all. Let's promote to some different notation now. Let's say I have a vector

$$v = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad (74)$$

There are four degrees of freedom here. Now, let's say that my vector must satisfy

$$v^\mu v_\mu = 5 \quad (75)$$

This is simply an equation, $a^2 + b^2 + c^2 + d^2 = 5$. This, like before, reduces our degrees of freedom by one. So, vector identities can be used in the same way as adding linear equations to a system of linear equations; they can eliminate degrees of freedom. This is done all the time in general relativity and cosmology. Additionally, we can work with tensors. Imagine a tensor that we can imagine as a 3x3 matrix

$$m_{\alpha\beta} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (76)$$

This has nine degrees of freedom. Now, if we said that our matrix was symmetrical $m_{\alpha\beta} = m_{\beta\alpha}$,

$$m_{\alpha\beta,\text{sym}} = \begin{bmatrix} a & b & c \\ b & e & f \\ c & f & i \end{bmatrix} \quad (77)$$

So there are six degrees of freedom. Now, if we decided that it was antisymmetric $m_{\alpha\beta} = -m_{\beta\alpha}$, you can never have non-zero diagonals

$$m_{\alpha\beta,\text{asym}} = \begin{bmatrix} 0 & b & c \\ b & 0 & f \\ c & f & 0 \end{bmatrix} \quad (78)$$

So there's only three degrees of freedom, which is quite constraining. Fortunately, we get to abuse this often.

Consider a tensor describable by a 4x4 matrix, which is most from general relativity. It would have 16, 10, and 6 degrees of freedom for the normal, symmetric, and antisymmetric cases respectively.

4.2 What the hell is an action?

The action (and the equations of motion) may be one of the most useful and universal pieces of physics that one can learn. With a good enough handle on how it works, you can solve any mechanics problem that you come across, as well as being able to entirely derive the laws of electromagnetism and gravity from first principles. Although we've all heard of "extremising the action" or something like it, this is something often obscured and left not fully understood. Let's see if we can fix that.

Imagine that you're moving a particle from point A to point B , like throwing a baseball. Which path does the particle take? In general, if I throw a baseball multiple times in the exact same manner, it appears to always take the same path, so somehow the baseball knows which path to take. How does it know?

You might be yelling into the page right now "well Ben, of course it knows where to go on a particular path, it just follows gravity. It will move along the steepest potential difference!"

First, there's no need to yell. Second, how does the baseball *know* to travel along the potential difference? Also, if the baseball has some momentum (like, from me throwing it), it won't travel perfectly along the potential gradient. Clearly the baseball is deciding how much to follow the potential difference at a given time, and how much to just continue on doing its thing⁴.

Let's think more about the baseball's motion. If the baseball is moving against the potential gradient, it would be quite unhappy about this. You'd have to push it to do so. But, just moving directly with the potential gradient also doesn't entirely make it happy, since it also wants to continue the direction in which it's already moving.

One could imagine the baseball assigning a "score" to every possible path between points A and B depending on how happy it would be to take it. This score is referred to as the *action* S . The baseball will take the path which minimizes this action. Since this score is decided over the whole length of the path, one can define the action as an integral over such a path

$$S = \int_A^B L dt \quad (79)$$

⁴I believe physicists call this inertia

Where we integrate from A to B over some time t ,⁵ over some unknown function L defined on the path. This function is called the *Lagrangian*. Since this is a function of our path, it will be a function of the *dynamical variables* in our system. For a system of particles, the dynamical variables are the positions and velocities of all the particles. For a system like a pendulum, it will be dependant on the angle of the pendulum. We write our dynamical variables as q_i and their time derivatives as \dot{q}_i . For now, let's pretend there's just one coordinate and drop the indexing. Since, in general, this could also be dependant on time, we can write our action as

$$S = \int_A^B L(q, \dot{q}, t) dt \quad (80)$$

Now, we can ask the particle what path it would like to go on. The particle tells us that it wants to take the minimum action⁶ and so we should look at when the variation of the action is very small. So, we can apply a variation to the action and multivariable-calculus it through the integral. Since this isn't time dependant, we only pick up the q and \dot{q} terms.

$$\delta S = 0 = \int_A^B \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) dt \quad (81)$$

Now, that second term looks ripe for integrating by parts

$$\int_A^B \frac{\partial L}{\partial \dot{q}} \delta \dot{q} dt = \left[\frac{\partial L}{\partial \dot{q}} \delta q \right]_A^B - \int_A^B \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \delta q dt \quad (82)$$

Now, we've been varying the path but each time we've been holding the beginning and end points fixed (I've told my baseball where it's starting from and where it's landing, and it's just choosing which way to go about it), and the first term is asking for δq at those boundaries. By construction, this is always zero, so we can ignore this term! We can substitute the above result into our varied action

$$\delta S = 0 = \int_A^B \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right) \delta q dt \quad (83)$$

For the integral to be identically zero, we can ask the integrand to be zero. This gives us

$$\frac{\partial L}{\partial q} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \quad (84)$$

Which are called the *Euler-Lagrange* equations. These give you the *equations of motion*, in other words, they tell you how your system evolves. In other other words, they tell the particle which path to travel on.

Now, in general, the Lagrangian is simply the kinetic minus potential energy of your system. It's beyond this document for now to prove this, but with it we can take an example. For a pendulum of length l , the only coordinate we care about is the pendulum angle θ . The potential energy $V = mgh = mgl(1 - \cos(\theta))$ is just based on the height of the pendulum. The kinetic energy $T = \frac{1}{2}mv^2 = \frac{1}{2}ml^2\dot{\theta}^2$ can be determined from calculating a circular velocity. So, the lagrangian can be written as

$$L = T - V = \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos(\theta)) \quad (85)$$

And we can apply the Euler-Lagrange equations, where $q = \theta$,

$$\frac{\partial}{\partial \theta} \left(\frac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos(\theta)) \right) = \frac{d}{dt} \left(\frac{\partial}{\partial \dot{\theta}} \left(\frac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos(\theta)) \right) \right) \quad (86)$$

Nicely, each derivative only hits one side,

$$-mgl \sin(\theta) = ml^2\ddot{\theta} \quad (87)$$

or,

$$\ddot{\theta} = -\frac{g}{l} \sin(\theta) \quad (88)$$

Which hopefully you remember from first year undergraduate physics. You can see how powerful and simple this is, and for more complicated systems this becomes the only feasible way to get the equations of motion.

And, for more complicated setups, you can write down Lagrangians for electromagnetism and gravity, which let you calculate their equations of motion. That is super powerful.

⁵In general this could be any parametrising variable. For classical mechanics, it's almost always time.

⁶In general, it's just an extremum of the action, so it could be a maximum or a saddle. In general relativity, it's actually usually the maximum.

Now, let's think about our Euler-Lagrange equations

$$\frac{\partial L}{\partial q} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \quad (89)$$

Notice how if we have a term in our Lagrangian that is constant relative to every q and \dot{q} , it will never appear in the equations of motion. More generally, you can add anything that is a total time derivative $L \rightarrow L + \frac{d}{dt} F(q, t)$, and it will drop out of every equation of motion. We can prove this by recalling the definition of the action,

$$S = \int_A^B L dt \quad (90)$$

Which, with the addition of the total time derivative, becomes,

$$S = \int_A^B L dt + \int_A^B \frac{d}{dt} F(q, t) dt \quad (91)$$

This new term can be easily solved

$$S = \int_A^B L dt + \left[F(q, t) \right]_A^B \quad (92)$$

The variation of this term becomes

$$\delta S = \delta \int_A^B L dt + \left[\frac{\partial F(q, t)}{\partial q} \delta q \right]_A^B \quad (93)$$

Which we remember from the derivation of the Euler-Lagrange equations is identically zero because we hold our boundaries fixed during the variation. So, this term does not appear in any equations of motion.

This is powerful, remember it. When you're confused about gauges, think back to here, they sort of work similarly to this.

4.3 Noether's Theorem

We can talk about conserved quantities in this Lagrangian formalism. One can find that each of these conserved quantities has a related symmetry, and this is called *Noether's Theorem*, which is my favourite theorem in physics.

Let's say that I change my coordinates through a particular transformation $q_i \rightarrow q_i + \delta q_i$. When we vary the coordinates in this way, the Lagrangian varies in this way as well.

$$\delta L = \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \quad (94)$$

Which we can rewrite by applying the Euler-Lagrange equation to the first term,

$$\delta L = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \quad (95)$$

We can see the first term has a time derivative hitting $\frac{\partial L}{\partial \dot{q}_i}$, while the second term has a time derivative hitting δq_i . This is because these two terms can be expressed as a derivative of a product, which means we can rewrite this equation as

$$\delta L = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) \quad (96)$$

Now let us imagine that the transformation that we've done is a *symmetry*. What this means is that changing the coordinates in this manner doesn't effect the physics of the system - so doesn't effect the Lagrangian, so $\delta L \equiv 0$,

$$0 = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) \quad (97)$$

Meaning the quantity inside the brackets is conserved over time! So, we've proved it, a symmetry generates a conserved quantity. You can be sure for the rest of your physics career to look for symmetries to find your conserved quantities.

But this isn't all, you could have your lagrangian being changed by this coordinate transform as being some additive term

$$L' = L + f(q, \dot{q}, t) \quad (98)$$

This doesn't happen all the time, but in some situations it does and we need to know how to handle it. Which means that a variation of the lagrangian needs to also satisfy

$$\delta L = f \quad (99)$$

Which means

$$f = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right) \quad (100)$$

or, if f happens to take the functional form of a total time derivative $f = \frac{d}{dt} F$,

$$0 = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i - F \right) \quad (101)$$

Which is the general statement of Noether's theorem. The conserved quantity is therefore

$$\frac{\partial L}{\partial \dot{q}_i} \delta q_i - F \quad (102)$$

For completeness, recall that the repeated use of the index i implies summation.

Let's build some intuition with the orbit example. Let's say I have a particle 2 dimensional space with a central potential $V(r)$. The Lagrangian will be, in polar coordinates,

$$L = T - V = \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\theta}^2 - V(r) \quad (103)$$

We can look at the equation of motion for θ , which through the Euler-Lagrange equations give

$$\frac{\partial L}{\partial \theta} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) \quad (104)$$

$$\implies \frac{d}{dt} (m r^2 \dot{\theta}) = 0 \quad (105)$$

Which is easy to evaluate since the Lagrangian is not explicitly dependant on θ . This tells you that $m r^2 \dot{\theta}$ is conserved, and this is actually the angular momentum. This is a result of the symmetry around θ : rotating an orbit on the plane doesn't change the physics of what the objects are doing. This is Noether's theorem in action: the symmetry around θ gives you a conserved quantity $m r^2 \dot{\theta}$. This can be seen more directly with the formula for Noether's theorem that we developed earlier.

Changing θ in the above system doesn't change the Lagrangian, so $F \equiv 0$. Then, we need to evaluate

$$\frac{\partial L}{\partial \dot{q}_i} \delta q_i \quad (106)$$

The confusing part is δq_i . This is just a variation, which can be written as a derivative, which in this case for a change in θ is a derivative with respect to θ .

$$\frac{\partial L}{\partial \dot{q}_i} \frac{dq_i}{d\theta} \quad (107)$$

You can read off that this θ derivative is 0 when $q_i = r$, and 1 when $q_i = \theta$. So,

$$\text{conserved quantity} = \frac{\partial L}{\partial \dot{\theta}} = m r^2 \dot{\theta} \quad (108)$$

Which is the same thing we had before.

4.4 Intuition From Electromagnetism (HEL 6)

Now we take a turn to motivating the field equations for electromagnetism to get a feel for how field theories work. We will move from our early undergraduate understanding of the maxwell equations, to a theory with tensors and indices and stuff like that. It makes it harder to look at, but more compact to write, and you will need to understand this notation to understand arbitrary fields, which we'll have to write this way.

We can start with the Lorentz force law for a particle with charge q moving with velocity \vec{u} through an electromagnetic field

$$\vec{f} = q(\vec{E} + \vec{u} \times \vec{B}) \quad (109)$$

Everything in this bracket can be written more generally as our velocity \vec{u} and some more complicated object \vec{F} that encodes everything captured in this equation.

$$\vec{f} = q\vec{F} \cdot \vec{u} \quad (110)$$

We can drop the vector nonsense and move into index forms, assuming that \vec{F} is a two-index tensor.

$$f_\mu = qF_{\mu\nu}u^\nu \quad (111)$$

Now, for interests sake, the 4-vector dot product $f_\mu u^\mu \equiv 0$ because it's orthorgonal to the force (this is because the 4-velocity is along the worldline, while the force is perpendicular to it, but this isn't important to know right now. Just know that this is using a 4-vector so there's also the time components), We can use this to write

$$f_\mu u^\mu = qF_{\mu\nu}u^\mu u^\nu = 0 \quad (112)$$

The only way for this to be true is for F to be antisymmetric when swapping around its two indices. Like mentioned a few sections ago, this will massively reduce the numbers of degrees of freedom.

We can now relate F to the source charges. In general, we describe this as F changing as a result of the source, which we'll call the current \vec{j} . In general, this 4-vector takes the form (ρ, I_x, I_y, I_z) where the time component is the charge distribution, and the spacial components describe the classical current. We write the field strength in terms of this general current,

$$\nabla \vec{F} \propto \vec{j} \quad (113)$$

We'll just say for now that they're equal, for our result this isn't hugely important, and we'll also switch to using indices

$$\partial_\mu F^{\mu\nu} = j^\nu \quad (114)$$

And, if you expand these out for the definition of F ,⁷ you recover Maxwell's equations as you know them.

An interesting thing to know is what happens as j^ν changes, so let's take a derivative of each side

$$\partial_\mu \partial_\nu F^{\mu\nu} = \partial_\nu j^\nu \quad (115)$$

But, as a result of the antisymmetry in F and commutability of partial derivatives, the left hand side will be identically zero. As a result,

$$\partial_\nu j^\nu = 0 \quad (116)$$

Which we call conservation of charge! There it is, from first principles.

Now, originally we said there's six degrees of freedom in F due to antisymmetry. But, our source charges generate only four equations (one for each index of ν), so the form of our tensor F must be more constraining. Crucially, in four dimensions a vector has four degrees of freedom. Clearly, for there to be only four degrees of freedom, F must be able to be written in terms of a vector A . One can write a tensor that is dependant on a single vector A like

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (117)$$

Which is a little bit arbitrary, but it is the convention. There are other ways one could do this, but this is the nicest and is agreed upon. A is known as the *Vector Potential*.

Let's now imagine we take our vector potential A and add some arbitrary function Q to it

⁷Which I won't do here, but basically you can represent it as a matrix where each entry is some component of \vec{E} or \vec{B} .

$$A'_\mu = A_\mu + Q_\mu \quad (118)$$

How would F change?

$$F'_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\mu Q_\nu - \partial_\nu Q_\mu \quad (119)$$

Now, $F' = F$, or in other words it is left invariant, if

$$\partial_\mu Q_\nu = \partial_\nu Q_\mu \quad (120)$$

Which occurs if Q is a derivative of some other function $Q_\mu = \partial_\mu \Psi$, because partial derivatives commute.

So, you can add things to the scalar potential that are a total derivative of this form, and they won't change the field tensor, and therefore don't change the equations of motion, and therefore don't change the physics. This is very similar to how total derivatives added to the Lagrangian leaves the equations of motion invariant. This is called *Gauge Invariance*, and abusing it is going to let us do many things in general relativity.

What this also does is constrain the structure of our field equations further. Since there's an infinite set of functions Ψ that we can add, this removes a degree of freedom from the field theory.

We can rewrite our field equations as, by substituting the definition of F in terms of A

$$\partial_\mu \partial^\mu A_\nu - \partial_\nu \partial^\mu A_\mu = j_\nu \quad (121)$$

The importance of the gauge symmetry is we can just choose a gauge wherein $\partial^\mu A_\mu = 0$.⁸ This relates all the values of A_μ to each other, which adds an additional constraint that removes an additional degree of freedom. So, in the end, we have only two degrees of freedom. You already know this for sure, but you've seen it called the two polarization states of light, or light being described by two superimposed transverse waves. We can demonstrate the second bit by making use of $\partial^\mu A_\mu = 0$ in our above equation

$$\partial_\mu \partial^\mu A_\nu = j_\nu \quad (122)$$

Introducing the *d'Alembertian* operator $\square^2 = \partial_\mu \partial^\mu$,

$$\square^2 A_\nu = j_\nu \quad (123)$$

$$\square^2 F_{\mu\nu} = 0 \quad (124)$$

Let's imagine one dimensional space, with time, and we explicitly write out the d'Alembertian operator with derivatives, the first equation looks like

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) A_\mu = j_\mu \quad (125)$$

The solution of this two superimposed plane waves (one real, one imaginary) travelling at the speed of light... exactly how we understand electromagnetism to work!

We can understand why there's two plane waves intuitively. Light is moving along the x direction. Special relativity prevents any wiggling of the photon along the x - and t -axes, since this would change lightspeed. So, the photon can only wiggle along the perpendicular y and z axis, which are represented by the two superimposed plane waves.

We had four degrees of freedom, but two were eliminated by the gauges. We call light a massless spin-1 particle. For a massive spin-1 particle, like a hypothetical graviton, it would be able to wiggle in the propagating direction as well (in other words, that $\partial^\mu A_\mu = 0$ constraint wouldn't apply) and so it would have three propagating degrees of freedom.

This example will help us work in more unintuitive field theories

5 Gravitational Waves (HEL 17&18)

I actually think we can build some more intuition here by working out the degrees of freedom in gravitational waves, and prove that they apply a transverse "stretchy-squeazy" ripple as they propagate. This will also be insightful for understanding perturbations.

⁸The fact that this works is left as an exercise.

5.1 Perturbations

Imagine we had some space where our metric $g_{\mu\nu}$ was only a little bit different than that of flat space $\eta_{\mu\nu}$

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad (126)$$

Here, $h_{\mu\nu}$ is a small tensor that describes our perturbations. Going forward, we can neglect terms of higher order in perturbation, since they will not contribute significantly to the physics. We can now do the classic move of considering a coordinate transform, and see what happens. Take, for example a transformation of the form

$$\tilde{x}^\mu = x^\mu + \xi^\mu(x) \quad (127)$$

And we are deciding that ξ^μ are small just like $h_{\mu\nu}$. We can easily work out that the jacobian will look like

$$\frac{\partial \tilde{x}^\mu}{\partial x^\nu} = \delta_\nu^\mu + \partial_\nu \xi^\mu \quad (128)$$

$$\frac{\partial x^\mu}{\partial \tilde{x}^\nu} = \delta_\nu^\mu - \partial_\nu \xi^\mu \quad (129)$$

Which we can use to work out how the metric changes. Recall that

$$\tilde{g}_{\mu\nu} = g_{\rho\sigma} \frac{dx^\rho}{d\tilde{x}^\mu} \frac{dx^\sigma}{d\tilde{x}^\nu} \quad (130)$$

$$\tilde{g}_{\mu\nu} = (\eta_{\rho\sigma} + h_{\rho\sigma})(\delta_\mu^\rho - \partial_\mu \xi^\rho)(\delta_\nu^\sigma - \partial_\nu \xi^\sigma) \quad (131)$$

Now recall that we can neglect higher order terms. So for example, when we multiply x^μ and $h_{\mu\nu}$ together, we can set that term to zero. Armed with this, we can obtain

$$\tilde{g}_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu \quad (132)$$

Which looks like our original equation for $g_{\mu\nu}$, except our ‘new’ $\tilde{h}_{\mu\nu}$ goes with $\tilde{h}_{\mu\nu} = h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu$. We can now see a connection with electromagnetism. In terms of the physics, if a transform of this form was made, the physics would remain unchanged. This is analogous to how rotating a Keplerian system doesn’t change physics, just the way you look at it. And, consequently, an associated conserved quantity exists. There’s a large freedom in allowed $h_{\mu\nu}$ that provide the same physics, and as a result, we lose degrees of freedom in the solution because of the generated conserved quantity! This is the *gauge transformation* and *Noether’s Theorem* coming into play to save us again.

5.2 Obtaining the field equations

The next objective is to see how this affects the field equations (which I have not yet derived.. for a future edition of my notes),

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -\kappa T_{\mu\nu} \quad (133)$$

Knowing that the Ricci tensor $R_{\mu\nu}$ is just a bunch of 2nd order derivatives of the metric, you can evaluate this. I’ll spare the algebra, since it’s not insightful. What is insightful is what we do with the result, so I’ll just write it down

$$\square^2 \bar{h}_{\mu\nu} + \eta_{\mu\nu} \partial_\rho \partial_\sigma \bar{h}^{\rho\sigma} - \partial_\nu \partial_\rho \bar{h}_\mu^\rho - \partial_\mu \partial_\rho \bar{h}_\nu^\rho = -2\kappa T_{\mu\nu} \quad (134)$$

For simplicity’s sake, $h_{\mu\nu}$ has been swapped out for the so-called ‘trace-reverse’ $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\text{Tr}(h_{\mu\nu})$. Henceforth we will be writing the trace of a tensor as simply the tensor’s character without any indices. In this situation, $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$.

This field equation is crazy complicated, but fortunately the gauges are here to save us. We can see how our coordinate transformations affect this trace-reverse tensor

$$\tilde{\bar{h}}_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\tilde{h} \quad (135)$$

$$= h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu - \frac{1}{2}\eta_{\mu\nu}(-\partial_\sigma \xi^\sigma - \partial_\rho \xi^\rho) \quad (136)$$

$$= \bar{h}_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu + \eta_{\mu\nu} \partial_\sigma \xi^\sigma \quad (137)$$

Similar to electromagnetism, it is then of interest to see how this transformed quantity $\tilde{\tilde{h}}_{\mu\nu}$ changes, which we can do by applying a derivative. First, we will switch to upper indices

$$\partial_\nu \tilde{\tilde{h}}^{\mu\nu} = \partial_\nu \bar{h}^{\mu\nu} - \partial_\nu \partial^\mu \xi^\nu - \partial_\nu \partial^\nu \xi^\mu + \partial_\nu \eta^{\mu\nu} \partial_\sigma \xi^\sigma \quad (138)$$

We can see that the last term is just a scalar, so the derivative will be zero. Likewise the mixed derivatives will make the second term drop out. This leaves us with

$$\partial_\nu \tilde{\tilde{h}}^{\mu\nu} = \partial_\nu \bar{h}^{\mu\nu} - \square^2 \xi^\mu \quad (139)$$

Here is where the gauge transformation works wonders. What if we choose our functions ξ^μ such that the right hand side of this equation is zero? Well, now $\partial_\nu \tilde{\tilde{h}}^{\mu\nu} = 0$. This is nice, and we can simply choose to work in this coordinate system, so we can drop the tilde. This is the power of the gauge transformation, recall how in covariant equations if you can satisfy an equation in one coordinate system, it's true in all coordinate systems? We're applying that here. We'll work in this convenient coordinate system and know that the physics works the same in the mathematically inconvenient ones.

This coordinate system is so convenient because it magically makes three of the terms in the field equation go away, and we only are left with

$$\square^2 \bar{h}^{\mu\nu} = -2\kappa T^{\mu\nu} \quad (140)$$

And we just need to recall that we require $\partial_\mu \bar{h}^{\mu\nu} = 0$.

But that's not all folks, we can also guarantee that $\partial_\nu \tilde{\tilde{h}}^{\mu\nu} = \partial_\nu \bar{h}^{\mu\nu}$ if $\square^2 \xi^\mu = 0$. This makes the quantity gauge invariant, and will be helpful shortly.

5.3 Stretchy-Squeezy

We can see that in the case of there being no stress-energy, we just get $\square^2 \bar{h}^{\mu\nu} = 0$, which is solved by two superimposed plane waves. The parallels with electromagnetism cannot be hitting you any harder.

But we can go further than this here. We can write out this solution explicitly,

$$\bar{h}^{\mu\nu} = A^{\mu\nu} e^{ik_\rho x^\rho} \quad (141)$$

Where $A^{\mu\nu}$ and k_ρ are some arbitrary functions called the *amplitude tensor* and *wavevector* respectively, which we can now constrain. First, we had the additional condition that $\partial_\mu \bar{h}^{\mu\nu} = 0$, which propagating this through our explicit plane-wave solution gives us

$$A^{\mu\nu} k_\mu = 0 \quad (142)$$

Now if we wanted to imagine a wave propagating in the x_3 direction (and also in time, of course), the wavevector would be $\vec{k} = (k, 0, 0, k)$ since there is no x_1 or x_2 components. From the above relation, $A^{3\nu} = A^{0\nu}$. And, since this is a symmetric tensor (which you can see from the fact that $h_{\mu\nu}$ has to be symmetric), we are reduced to six degrees of freedom. However, we can then apply the additional $\square^2 \xi^\mu = 0$ gauge transformation. This can be written as a solution of plane waves like earlier,

$$\xi^\mu = \epsilon^\mu e^{ik_\rho x^\rho} \quad (143)$$

Then, we can see how the amplitude tensor transforms,

$$\tilde{A}^{\mu\nu} = A^{\mu\nu} - i\epsilon^\mu k^\nu - i\epsilon^\nu k^\mu + i\eta^{\mu\nu} \epsilon^\rho k_\rho \quad (144)$$

Here's the fun part, we can choose our functions ϵ^μ such that there are only two degrees of freedom remaining. We can choose that $\tilde{A}^{0\nu} = 0$, $\tilde{A}^{11} = -\tilde{A}^{22} = a$, and $\tilde{A}^{12} = b$. This, in technical terms, is called the *transverse-traceless gauge*. We can write the amplitude tensor as

$$A_{TT}^{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & b & -a & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (145)$$

Where now the values a and b describe our two possible polarization states. At this point, the comparisons with electromagnetism should be drilled into your mind. Now, on top of considering our wave propagating in the x_3 direction, consider it to have $b = 0$ polarization. We'll now write our amplitude tensor with $A_{TT}^{\mu\nu} = a e_a^{\mu\nu}$, where e_a is the *polarization tensor*. It's just our amplitude tensor but with only 1s and 0s. We can now write out our metric perturbation, which pulls out the real part of our plane waves

$$h_{TT}^{\mu\nu} = ae_a^{\mu\nu} \cos(k_\rho x^\rho) \quad (146)$$

Recall that we're propagating along the x_3 direction with a wavevector $\vec{k} = (k, 0, 0, k)$. This gives us

$$h_{TT}^{\mu\nu} = ae_a^{\mu\nu} \cos k(x_0 - x_3) \quad (147)$$

Or, if we just call $x_3 = x$ and rewrite x_0 as our time,

$$h_{TT}^{\mu\nu} = ae_a^{\mu\nu} \cos k(ct - x) \quad (148)$$

Fantastic, this just shows our wave travelling through the space at the speed of light (Do I have to remind you just how similar this is to electromagnetism? Have I mentioned that?).

Well, here's the cool bit. We can describe the distance between two points in space with our line element. Here we will use a vector ξ to represent the orientation of the two nearby points, and we use subscripts ij to correspond to only the spacelike coordinates (we're not in four dimensional spacetime here, we don't care about that).

$$D^2 = -g_{ij}\xi^i\xi^j = (\delta_{ij} - h_{ij})\xi^i\xi^j \quad (149)$$

But this is the coordinate distance, what about physical, actual distance? What about writing this so that g_{ij} is incorporated into the coordinates? After all, some change in stress-energy will change g_{ij} and if we instead want to see how our points move around compared to flat space, we should push that into the coordinates themselves. This is like moving from normal to conformal time in cosmology, if you've seen that before. We can do this by introducing a new coordinate ζ^i ,

$$\zeta^i = \xi^i + \frac{1}{2}h_k^i\xi^i \quad (150)$$

You can confirm yourself that plugging this into the distance formula gives,

$$D^2 = \delta_{ij}\zeta^i\zeta^j \quad (151)$$

Now, we can see how our physical coordinate changes!

$$\zeta^i = \xi^i + \frac{1}{2}h_k^i\xi^i = \xi^i + \frac{1}{2}ae_{j,a}^i \cos k(ct - x)\xi^j \quad (152)$$

$$\downarrow \quad (153)$$

$$\zeta^1 = \xi^1 - \frac{1}{2}a \cos k(ct - x)\xi^1 \quad (154)$$

$$\zeta^2 = \xi^2 + \frac{1}{2}a \cos k(ct - x)\xi^2 \quad (155)$$

$$\zeta^3 = \xi^3 \quad (156)$$

And there it is! What does this mean? the 3rd component of ζ doesn't change, but the first and second components oscillate. So, as a gravitational waves propagates along the x_3 direction, x_1 and x_2 oscillate stretching and squeezing!