PHY 387M: Relativity Theory

Delon Shen

Notes for Prof. Matzner's Relativity Theory(PHY 387M) course at UT Austin during Spring 2021. The course follows Misner, Thorne, and Wheeler's "Gravitation" as well as Prof. Matzner's own notes. This will also contain my reading notes from some parts of Wald's *General Relativity*. If you have any comments let me know at hi@delonshen.com.

Last updated on February 10, 2021

Lecture 1a: Historical Background and Special Relativity

January 17, 2021

Historical background I'm leaving out goes here

Let $\mathcal{E}$ be an event in a $D = 4$ spacetime $\mathcal{M}$. $\mathcal{E}$ could be a camera flash going off at position $x^\mu = \{t, x, y, z\}$. Lets say we have two such events $\mathcal{E}_1$ and $\mathcal{E}_2$. The interval between these two events is

$$x^\mu x_\mu = x^\mu g_{\mu\nu} x^\nu = -c^2 t^2 + x^2 + y^2 + z^2$$

This interval is the same in any reference frame. Lets try to derive the Lorentz transform. Consider some fast guy walking towards you with speed v. You're standing still in your reference frame. Your position in the fast guy's reference frame x' if we assume Galilean relativity is $x = x' - vt$. However this clearly doesn't keep the speed of light c the same in every reference frame. Thus lets introduce an undetermined function $\gamma(|v|)$ where we use $|v|$ to impose isotropy

$$x' = \gamma(|v|)\left(x - \frac{v}{c}ct\right) \tag{1}$$

In Galilean relativity $ct' = ct$ but this also doesn't work. However, if we're somehow inspired to, we can also guess for special relativity

$$ct' = \gamma(|v|)\left(ct - \frac{v}{c}x\right) \tag{2}$$

Now we use the invariant interval to get γ

$$-c^2t^2 + x^2 = -\gamma^2\left(ct - \frac{v}{c}x\right)^2 + \gamma^2\left(x - \frac{v}{c}ct\right)^2$$

Solving for γ with Mathematica gives us the following

$$\boxed{\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}}$$

We should also know that the invariant interval can become infinitesimal giving us an infinitesimal arc length in flat space time

$$-ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$$

This leads us to define the four velocity

$$\frac{dx^\mu}{ds} = \left\{c\frac{dt}{ds}, \frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds}\right\}$$

We define spacelike as $ds^2 > 0$, timelike as $ds^2 < 0$, and null (e.g. light ray) as $ds^2 = 0$. Note that if $ds^2 = 0$ we can't define the 4-velocity as above. We introduce some parameter (affine parameter?) λ and have

$$0 = -c^2\left(\frac{dt}{d\lambda}\right)^2 + \left(\frac{dx}{d\lambda}\right)^2 + \ldots$$

From now on we will let $c = 1$.

Lecture 1b: Some examples in special relativity

January 19, 2021

Lightcone stuff here

We'll now introduce the metric $\eta_{\mu\nu}$ (for some reason he uses opposite signature as what he did last lecture?)

$$\eta_{\mu\nu} = \text{diag}(-1,1,1,1) \Rightarrow ds^2 = \eta_{\alpha\beta}dx^\alpha dx^\beta = -dt^2 + dx^2 + dy^2 + dz^2$$

Lets look at time dilation. Say we're standing still in our reference frame K. Now consider a moving frame K' with velocity v. In the K' frame we know that $d\tau' = dt'$. From (1) (2) we know that

$$dx' = 0 = \gamma(dx - vdt) \Rightarrow dx = vdt \Rightarrow dt' = \gamma(dt - v^2dt) = \gamma dt/\gamma^2 = dt/\gamma$$

The last equality is time dilation. Now for length contraction. Consider a meter stick sitting at rest in reference frame K. Now consider an observer moving with velocity v with resepct to K. The moving observer's rest frame is K'. Now we measure the lenght of the meter stick in the K' frame as l', (in this case $dt' = 0$). Now again with what we found in lecture 1a

$$dt' = 0 = \gamma(dt - vl) \Rightarrow dt = vl \Rightarrow l' = \gamma(l - (vdt = v^2l)) = \gamma(1 - v^2) = l\gamma/\gamma^2 \Rightarrow l' = l/\gamma$$

Here we have length contraction.

Skipping einstein summation stuff since QFT has beaten that into me

Twins Paradox

I got the material in this subsection from here as well as Matzner's lecture on this stuff. A and B are a couple who happen to be born at exactly the same time. B is going on a space mission. He will get on a rocket ship and travel away from earth at a velocity V for some time T and then will travel back to earth with velocity $-V$ for the same time T. Thus A will have aged $2T$ in the time that B has been gone but from time dilation he expects B to be younger than he is when B gets back. However by symmetry B would expect A to be younger when B gets back since from B's perspective, A is travelling away from him. Lets resolve this paradox. First we'll formalize what we said above by defining a few events. Let a_1 be the event when B leaves earth, a_2 be the event when B turns around, and a_3 be the event when B returns to earth. The proper time elapsed for A from a_1 to a_2 is $\tau_A(a_1 \to a_2) = T$ and similarly $\tau_A(a_2 \to a_3) = T$. This gives us $\tau_A(a_1 \to a_3) = 2T$. From our result on time dilation above we get that $\tau_B(a_1 \to a_3) = 2T\sqrt{1 - V^2}$. Now to second order in V (we could go to higher order but the first non-vanishing term in the Taylor expansion illustrates what we'll want to get across)

$$\tau_A - \tau_B = 2T(1 - \sqrt{1 - V^2}) \approx TV^2 + O(V^3)$$

A is older than B. But in B's reference frame we'd expect B to be older than A by symmetry. There a $2TV^2$ term missing somewhere that points to an asymmetry. So where does the asymmetry come in? Lets look at a_2 more closely. Lets assume B accelerates backwards with acceleration g for some $\delta t'$ where $\delta t' \ll T$. We know that

$$g\delta t' = 2V$$

Now note that to first order in V from (1) (2) we have

$$x' = \gamma(x - Vt) \approx x - tV + O(V^2) \Rightarrow x \approx x' + tV$$

$$t' = \gamma(t - Vx) \approx \gamma(t - V(x' + tV)) \approx t - x'V \Rightarrow t = t' + x'V$$

If we want to assert acceleration we'll let $V = g\delta t'$ meaning that

$$\delta t = \delta t'(1 + gx') \Rightarrow \frac{\delta t}{\delta t'} = 1 + gx'$$

What this equation is saying is that in an accelerating frame at different "height" (e.g. x' which is TV in this case), A is aging at a different rate than B. This will resolve our paradox. Since we're accelerating to the left the acceleration is $-g$ and the position of A in B's frame is $-TV$ meaning that

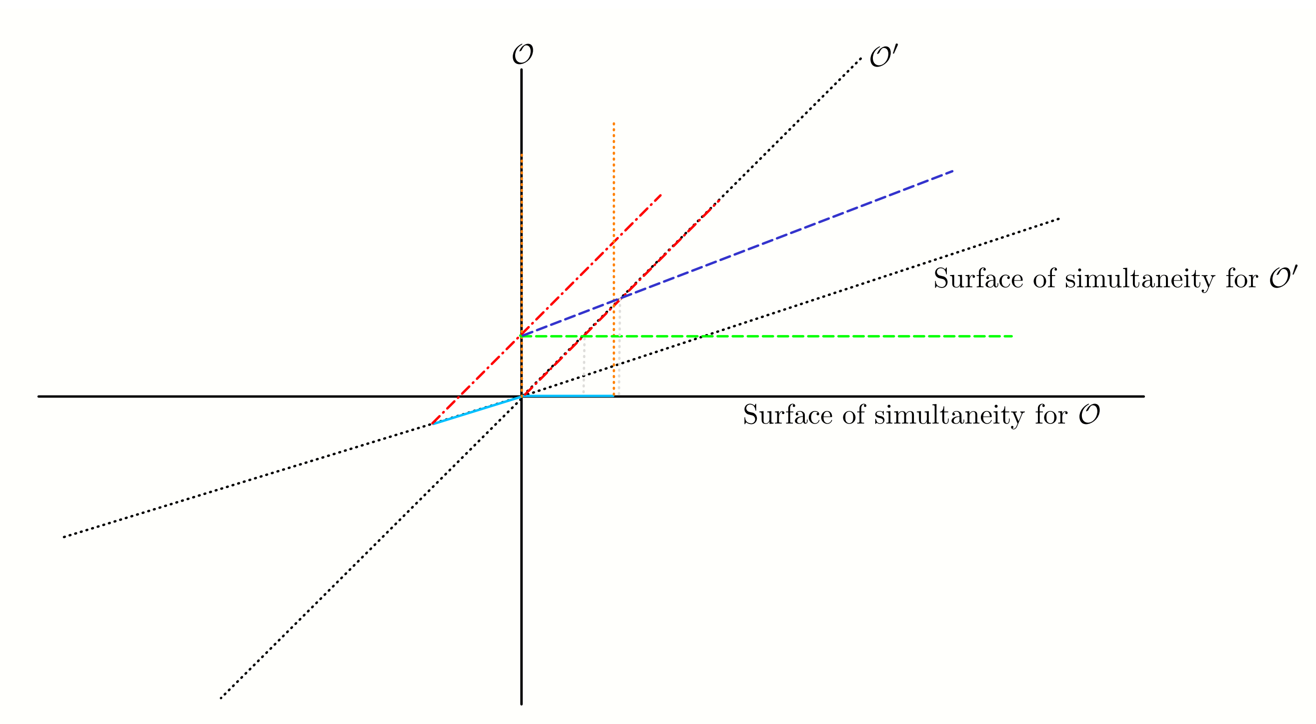
$$\delta t - \delta t' = \delta t' g x' = 2TV^2$$

The missing $2TV^2$ term that resolves the idea that A should be older than B by TV^2 if we're following along with B. This difference in passage of time at different heights in an accelerating frame can also be measured by GPS's (I think Matzner mentioned this in the lecture.) One thing to note is that we only resolved the paradox to order V^2 but it gets the point across and is valid for higher orders according to Hirata (I'll just take his word on this here.)

WALD PROBLEM 1.1: CAR AND GARAGE PARADOX

January 19, 2021

Taking inspiration from figure 1.3 of Wald we get



Where O is the observer at rest in the garage and O' is the moving vehicle observer frame. From the dashed green line we see that for $\mathcal{O}$ when the back of the car enters the garage the front of the car is still in the garage thus the doorman is correct in his frame. From the dashed purple line we see that for $\mathcal{O}'$ when the back of the car enters the garage the front of the car has already gone through the back of the garage and thus $\mathcal{O}'$ is also correct in his assumption.

Wald Chapter 2: Manifolds and Tensor Fields

Started: Started: January 20, 2021. Finished: Not for a While Cause I'm not done with problems :(

Lets start by motivating the idea of manifolds. Before general relativity we could assume that globally space time was flat, $\mathbb{R}^{3,1}$. However with the entrance of general relativity we'll be solving for the global structure of spacetime. Locally however we can still say things looks flat. Trying to solve for the structure of ST is similar to trying to determine the shape of the Earth as a sailor. Locally we know that the surface of Earth looks like $\mathbb{R}^2$ but globally it wouldn't be safe to assume that the surface of the Earth is $\mathbb{R}^2$. We do know however that the surface of the Earth is some surface embedded in $\mathbb{R}^3$ so this could motivate us to study ST as an embedding in some higher dimensional space. However ST doesn't have a natural higher dimensional space which we can embed it into.

So lets try to formalize this notion of a manifold. First we'll define a open ball int $\mathbb{R}^n$ of radius r centered at point $y = (y^1, \dots, y^n)$ as all point $x = (x^1, \dots, x^n)$ where

$$|x - y| = \left\{ \sum_{i=1}^{n} (x_1 - y_1)^2 \right\}^{1/2} < r$$

A open set is a set which is the union of a set of open balls. Now a manifold is sort of like a patchwork of open subsets of $\mathbb{R}^n$. Or more formally a real, smooth manifold $\mathcal{M}$ is a set and a set of subsets $\{O_\alpha\}$ that satisfy the following property

(a) $\{O_\alpha\}$ cover $\mathcal{M}$

(b) For every α there exists a bijective function ψ_α between O_α and U_α where U_α is some subset of $\mathbb{R}^n$

(c) If there exists α and β such that $O_\alpha \cap O_\beta \neq \emptyset$ then there exists a smooth function $\psi_\beta \circ \psi_\alpha^{-1}$ which takes points from $\psi_\alpha(O_\alpha \cap O_\beta) \subset U_\alpha \subset \mathbb{R}^n$ to points $\psi_\beta(O_\alpha \cap O_\beta) \subset U_\beta \subset \mathbb{R}^n$. We'll also require that these subsets U_α be open subsets.

The ψ_α are what we call coordinate systems. To make sure we can't create new manifolds by removing or adding coordinate systems we could also require $\{\psi_\alpha\}$ to contain all functions that satisfy (b) and (c) (e.g. maximal).

Wald Appendix A.1: Some Definitions for Topological Spaces

The reason we care about topological spaces is because general relativistic space time has the structure of a topological space (and more!) So lets define a topological space . Let X be a set and $\mathcal{J}$ be a collection of subsets of X. We define $(X, \mathcal{J})$ to be a topological space that satisfies the following properties

(a) Let $\{O_\alpha\} \subset \mathcal{J}$. We require $\bigcup_\alpha O_\alpha \subset \mathcal{J}$

(b) Let $\{O_\alpha\}$ be a finite subset of $\mathcal{J}$. We require $\bigcap_\alpha O_\alpha \subset \mathcal{J}$.

(c) We require $\emptyset$ and X to be members of $\mathcal{J}$.

$\mathcal{J}$ is a topology on X and contains only open sets. Lets look at some fun things we can do with topological spaces

(a) We can make a topology out of any set pretty easily. For example $\{X, \{X, \emptyset\}\}$ works pretty well. That's a fun party trick I guess.

(b) Let $X = \mathbb{R}$. We can let $\mathcal{J}$ to be the set of all sets which can be formed by the union of some open interval (a, b). This generalizes to $\mathbb{R}^n$ with open balls of $\mathbb{R}^n$.

(c) For a given topological space $\{X, \mathcal{J}\}$ any subset $A \subset X$ can also be made into a topology by definition of a topology of A as $\mathcal{I} = \{U | U = A \cap V \text{ s.t. } V \in \mathcal{J}\}$. This is called an *induced topology.*

(d) Let $\{X_1, \mathcal{J}_1\}$ and $\{X_2, \mathcal{J}_2\}$ be topological spaces. We can define a topology $\mathcal{J}$ for the set $X = \{(x_1, x_2) \text{ s.t. } x_1 \in X_1,\ x_2 \in X_2\}$ as all sets that are the unions of sets of the form $O_1 \times O_2$ where $O_i \in \mathcal{J}_i$. In this way we can build up to a topological space for $\mathbb{R}^n$

(e) Let $\{X, \mathcal{J}\}$ and $\{Y, \mathcal{J}\}$ be topological spaces. Consider a map $f : X \to Y$. If for any open subset $O \in J$, $f^{-1}(O)$ is open as well we call this map continuous. Now if f is continuous and bijective and f^{-1} is continuous we say that f is a *homeomorphism* and the two topological spaces are *homeomorphic*, they have the same topological properties.

(f) From here on out assume $\{X, \mathcal{J}\}$ is a topological space

(g) A set $C \subset X$ is closed if $X - C$ is open. In the topology we described in (a) we see that all members of $\mathcal{J}$ are both open and closed. We define topologies where the only subsets that are both open and closed are the set itself and $\emptyset$ as *connected* .

(h) Let $A \subset X$. The *closure* of A denoted by $\overline{A}$ is the intersection of all closed sets containing A. $\overline{A}$ contains A, is closed, and equal A iff A is closed as well. Similarly the *interior* of A denoted by $\tilde{A}$ is the union of all open sets inside A. $\tilde{A}$ is a subset of A, is open, and equal A iff A is open as well. The *boundary* of A is the set of point in $\overline{A}$ that are not in $\tilde{A}$.

(i) If $\forall p, q \in X$ we can find $O_p, O_q \in \mathcal{J}$ such that $p \in O_p$ and $q \in O_q$ then we call that topological space *Hausdorff*

(j) A set A is said to be compact if for any open cover of A there exists a finite subcover of A.

(k) Let $\{O_\alpha\}$ be a open cover of X. We call another open cover $\{V_\alpha\}$ a *refinement* of $\{O_\alpha\}$ if for all β there exists an O_α such that $V_\beta \subset O_\alpha$.

(l) We call a cover $\{V_\alpha\}$ locally finite if for all $x \in X$ there exists a neighborhood W such that the number of sets $V \in V_\beta$ that satisfy $V \cap W = \emptyset$ is finite.

(m) A topological space is paracompact if for each open cover $\{O_\alpha\}$ of X there exists a locally finite refinement of $\{O_\alpha\}$.

We'll stop with appendix A here. For chapter two we only need the definitions.

Back to Manifolds

If we take the topological route to define manifolds we would require the set of $\{\psi_\alpha\}$ to contain only homeomorphic functions. The only topological spaces considered in this book are *Hausdorff* and *paracompact.*

Lets consider an example of a manifold, a 2-sphere S^2, which we define as

$$\{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1^2 + x_2^2 + x_2^3 = 1\}$$

To map this onto $\mathbb{R}^2$ we define our elements of the covering set as $O_i^\pm$ such that

$$O_i^\pm = \{(x_1, x_2, x_3) \in S^2 \wedge \pm x_i > 0\}$$

Namely the set of $O_i^\pm$ is the set of six hemispheres that cover S^2. Also we can use a homomorphic function to project each $O_i^\pm$ onto $U_\alpha = D \subset \mathbb{R}^2$ where D is the disk on the j,k plane and $i \neq j \neq k$. This also satisfies the condition for overlapping elements of the cover of S^2. Namely $\psi_\beta \circ \psi_\alpha^{-1}$ behaves exactly as we expect it to. We prove this in the end of chapter problems

We can also define products of manifolds. Consider two manifolds M and M'. We can define $M \times M' = \{(p,p') | p \in M \ \wedge p' \in M'\}$. From here we can construct the covering set for $M \times M'$ by considering the covering set for M which we denote by O and M' which we denote by O'. We get $O_{\alpha\beta} = \{(O_\alpha, O'_\beta) | O_\alpha \in O \wedge O'_\beta \in O'\}$. Also for all α and β there should exists an $\psi_{\alpha\beta}$ such that $\psi_{\alpha\beta}(O_{\alpha\beta}) = U_{\alpha\beta}$. We can construct this as you would expect. By letting $\psi_{\alpha\beta}(\{O_\alpha, O'_\beta\}) = \{\psi_\alpha(O_\alpha), \psi'_\beta(O'_\beta)\}$ where ψ_α is the corresponding function for O_α in M and ψ'_β is the corresponding function for O'_β in M'.

We can now describe differentiability and smoothness. Consider manifolds M and M' with coordinate systems $\{\psi_\alpha\}$ and $\{\psi'_\alpha\}$ respectively. We call a map f smooth if for all α and β we have that $\psi'_\beta \circ f \circ \psi_\alpha^{-1}$ is a smooth function between U_α and U_β. Furthermore if this f is one-to-one, onto, and has a smooth inverse map then we call this function a *diffeomorphism* and the two manifolds *diffeomorphic*.

Vectors

We all know about vector spaces. You know about them, I know about them. However our intuitive notion of vector spaces start to break down in curved manifolds. For example, how do we define a vector space on a 2-sphere so that the vector space is still closed under vector addition? We'll find that we can recover our intuitive notion of vector spaces by considering *infinitesimal* vectors which stems from the fact that in general relativity we can assume that locally space looks flat (think about flat earthers.) However it turns out that our intuition for infinitesimal vectors breaks a little in curved geometry as well. For a sphere we have an intuitive picture of a tangent vector to the sphere since it's embedded in $\mathbb{R}^n$. However when we no longer have ourselves embedded in $\mathbb{R}^n$ our intuition for tangent spaces becomes a little shaky. So we'll start by trying to construct tangent spaces from only the properties of the manifold and keep $\mathbb{R}^n$ as our favorite special case for checking our work.

The way we'll construct this notion of tangent vectors is through direction derivatives along those tangent vectors. For a quick refresher on directional derivatives lets consider $\mathbb{R}^2$ say we want to find the change of a function $f(x,y)$ along a vector $\mathbf{v} = a\hat{x} + b\hat{x}$. Then we can write the direction derivative of f along $\mathbf{v}$ as $D_{\mathbf{v}}f$ with the definition of derivatives we learned in our first calculus class

$$D_{\mathbf{v}}f = \lim_{h\to 0} \frac{f(x+ah, y+bh) - f(x,y)}{h}$$

This so far isn't very illuminating but we can bring this into a cleaner form. Consider the function $g(z) = f(x_0 + az, y_0 + ba)$ where everything except for z is fixed. The derivative of this with respect to z can be again found with our regular derivative definition

$$g'(z) = \lim_{h\to 0} \frac{g(z+h) - g(z)}{h} \Rightarrow g'(0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h} = \lim_{h\to 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} = D_{\mathbf{v}}f(x_0, y_0)$$

Still not very illuminating. But now lets consider $g(z) = f(x = x_0 + az, y = y_0 + bz)$, the same function but dressed differently. Using the chain rule we get

$$g'(z) = \frac{\partial g}{\partial x}\frac{dx}{dz} + \frac{\partial g}{\partial y}\frac{dy}{dz} = \frac{\partial f}{\partial x}a + \frac{\partial f}{\partial y}b \Rightarrow g'(0) = \left(\frac{\partial f}{\partial x}a + \frac{\partial f}{\partial y}b\right)\bigg|_{x=x_0,\ y=y_0} = D_{\mathbf{v}}f(x_0, y_0)$$

Now since we fixed x_0, y_0 arbitrarily we now have a much more useful definition for directional derivatives

$$D_{\mathbf{v}}f(x,y) = v_x \frac{\partial f}{\partial x} + v_y \frac{\partial f}{\partial y}$$

This inspires the definition of direction derivatives which we'll use to implicitly define vectors and thus tangent vectors on our manifold.

$$v = v^{\mu}\partial_{\mu}$$

Let F be the set of smooth functions from M to $\mathbb{R}$. We define a tangent vector at a point $p \in M$ as the map $v : \mathsf{F} \to \mathbb{R}$. Note that this map is linear and obeys its own Leibnitz rule.

Theorem: Consider a n−dimensional manifold called M and some $p \in M$. Also let V_p denote the tangent space at p. We will show that $\dim V_p = n$.

Proof: We can do this by explicitly constructing an orthogonal basis for V_p with n elements. To do this first consider some coordinate system ψ and some function $f : U \to \mathbb{R}$. (I think) a concrete example of what f could be is the temperature at each point on the manifold. With these two function we can define a function

$$F = f \circ \psi^{-1} : \mathbb{R}^n \to \mathbb{R}$$

This is a function on the coordinates that label a manifold instead of the manifold itself. E.g. we can ask what's the temperature at $(r = R, \theta = 0, \phi = 0)$ with F wheras with f we could only ask what's the temperature at the pole. From here we will define X_{μ} which we'll show is our orthogonal basis for V_p

$$X_{\mu} = \frac{\partial}{\partial x^{\mu}} F\bigg|_{\psi(p)}$$

This function is determing the rate of change of f along the basis coordinates defined by coordinate system ψ(or something like that?). These are tangent vectors. Now we'll use a result

(TODO: it'll be proven in the end of chapter problems) that basically says for smooth functions g shift the origin.

$$g(x) = g(a) + (x_\mu - a_\mu)H^\mu(x)$$

I tried my best to illustrate this in Figure 1 Asserting that $g = F$ and $a = \psi(p)$ we get for a point

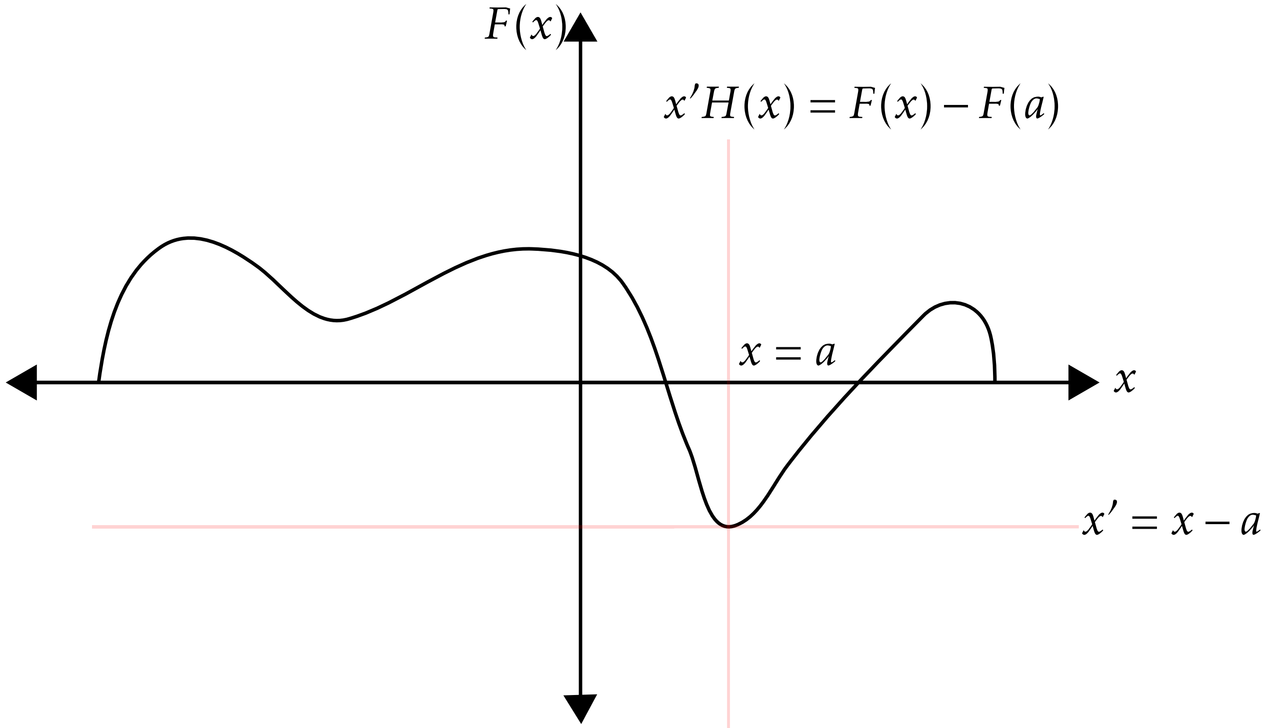


Figure 1: Visual depiction of the result of problem 2.2 for a one dimensional manifold. This was used to prove that the dimension of a tangent space V_p is n

q on the manifold

$$f(q) = f(p) + (x^\mu \circ \psi(q) - x^\mu \circ \psi(p))H_\mu(\psi(q)) \tag{3}$$

From considering an infinitesimal displacement $q = p + \delta p$ we can also intuit that

$$H_\mu(\psi(p)) = \frac{\partial F}{\partial x^\mu}\bigg|_{q=p} \tag{4}$$

Now lets consider a arbitrary tangent vector $v \in V_p$. What we want to do is find the directional derivative of this function $f(q)$ along this v at the point on the manifold p. To do this we apply the definition of $v = v^\mu \partial_\mu$ to Equation 3. We use the linearity and leibnitz rule for v as well as the fact that the directional derivative of a constant is zero.

$$v(f) = H_\mu(\psi(q))\bigg|_{q=p} v[x^\mu \circ \psi - \underline{x^\mu \circ \psi(p)}] + \underline{[x^\mu \circ \psi - x^\mu \circ \psi(p)]}\bigg|_p v[H_\mu \circ \psi]$$

Now applying Equation 4 to the first term and noticing that the underlined terms vanish we're left with

$$\boxed{v(f) = \frac{\partial F}{\partial x^\mu}\bigg|_{q=p} v[x^\mu \circ \psi] = X_\mu v[x^\mu \circ \psi]} \tag{5}$$

Now we can see that X_μ is a basis for V_p and the components of an arbitrary tangent vector $v \in V_p$ are determined by $v[x^\mu \circ \psi]$. What this is saying is that any tangent vector can be constructed by the directional derivative in the direction of x_μ of some function f times the component vector x_μ. Since we have proven that X_μ is a basis for V_p we have shown that dim $V_p = n$.

This X_μ is called the coordinate basis. If we had chosen another coordinate system ψ' then we could find the coordinate basis in the new coordinate system. First let $x^{\nu'}$ the ν component of $\psi' \circ \psi^{-1}$. We can then assert

$$X_\mu = \frac{\partial}{\partial x^\mu}\bigg|_{\psi(p)} = \frac{\partial x^{\nu'}}{\partial x^\mu}\frac{\partial}{\partial x^{\nu'}}\bigg|_{\psi(p)} = \frac{\partial x^{\nu'}}{\partial x^\mu} X_{\nu'}\bigg|_{\psi(p)}$$

Lets try plugging this into (5). We know geometrically v stays the same in both coorinate systems so we need to equate

$$X_\mu v^\mu = X_{\nu'} v^{\nu'} \Rightarrow \frac{\partial x^{\nu'}}{\partial x^\mu} X_{\nu'} v^\mu = X_{\nu'} v^{\nu'} \Rightarrow \boxed{\frac{\partial x^{\nu'}}{\partial x^\mu} v^\mu = v^{\nu'}} \tag{6}$$

The boxed equation is the *vector transformation law.*

Now lets consider an arbitary curve C on the manifold. We can think of C as a smooth map from $\mathbb{R}$ to the manifold. At each point p on this curve we can associate with it a tangent vector $T \in V_p$. Let $f \in \digamma$ be a function from the manifold to $\mathbb{R}$. By our definition of tangent vectors we have.

$$T(f) = \frac{\partial (f \circ C)}{\partial t}$$

Also note for an arbitrary coordinate system ψ we can write $\psi \circ C \Leftrightarrow x^\mu(t) \Rightarrow C = \psi^{-1}(x^\mu(t))$. Thus we can write

$$T(f) = \frac{\partial (f \circ \psi^{-1})}{\partial x^\mu}\frac{\partial x^\mu}{\partial t} = X_\mu \frac{\partial x^\mu}{\partial t}$$

From this we can see that the components of the tangent vector is given by

$$\boxed{T^\mu = \frac{\partial x^\mu}{\partial t}}$$

Definition $[v, w](f) = v[w(f)] - w[v(f)]$ skip formalism for fields for now. this part feels more intuitive than last part.

Tensors

First we will define a *dual vector* as a linear map that takes some number of spatial displacement vectors and maps them to a number. Or more formally consider a vector space V. Let V^* be the collection of linear maps $f : V \to \mathbb{R}$. Since V is a vector space it is easy to see that due to f being linear that V^* is also a vector space. Now let us define a basis $\{v^{\mu *}\}$ with the defining property that that the action of this basis on the basis of V is

$$v^{\mu *} v_\nu = \delta^\mu_\nu \tag{7}$$

We will show in the problems that this makes $\{v_{\mu *}\}$ a basis for V^*. From here we call V^* the dual space to V. Right now V doesn't have enough strucutre to uniquely define its dual space V^*. This extra strucutre will come from a metric. Now lets prove that the dual of the dual space V^* denoted by V^{**} can be identified to V. This part is based on these set of notes since Wald was too slick for me here. Let consider a map $g : V \to V^{**}$ defined as $g(v)(f) = f(v)$ where $v \in V$ and $f \in V^*$. To show that the two spaces can be identified with eachother we'll need an isomorphic function between the two spaces. Our canidate for this function is g. To show that g is isomorphic we'll need the following components.

(a) We know that since V and V^{**} have the same dimension we only need to prove that g is one-to-one to prove that g is isomorphic.

(b) To prove that g is one to one we need to prove that null g is the zero element.

Let $v \in \text{null}(g)$. Then $g(v)$ is the zero element of V^{**}. This means that for any $v^* \in V^*$ we have $(g(v))(v^*) = 0$. Thus by the definition of our function for any linear function $v^* \in v$ we have $v^*(v) = 0$. This property is only satisfied if $v = 0$. Thus null $g = \{0\}$ and thus g is one-to-one. From now on identify V with V^{**}.

Lets define a *tensor* T of type (k,l) as

$$T : \underbrace{V^* \times \cdots \times V^*}_{k} \times \underbrace{V \times \cdots \times V}_{l} \to \mathbb{R}$$

Namely we take in k dual vectors and l vectors and return a real number. For example a $(1,0)$ tensor takes in a dual vector and returns a number. Let $\mathcal{T}(k,l)$ be a vector space of all (k,l) tensors. A *contraction* $C : \mathcal{T}(k,l) \to \mathcal{T}(k-1,l-1)$ on the i^{th} dual vector and j^{th} vector is an operation that inserts $v^{\nu'}$ as the i^{th} dual vector and v_ν to the j^{th} vector. For example above we defined the $(1,1)$ tensor $v_{\mu *} v^\nu$. The contraction would then yield $v_\mu v^\mu = D$ where D is the dimension of the space we're considering (aka the trace of the kronecker delta). A *outer product* takes a tensor T of rank (k,l) and a tensor T' of rank (k',l') and define a tensor $T \otimes T'$, a rank $(k+k', l+l')$ tensor. If we have as input $\{v_1^*, \ldots, v_{k+k'}^*\}$ dual vectors and $\{v_1, \ldots, v_{l+l'}\}$ vectors then $T \otimes T'$ acts as

$$T(v_1^*, \ldots, v_k^*, v_1, \ldots, v_l) \times T'(v_{k+1}^*, \ldots v_{k+k'}^*, v_{l+1}, \ldots, v_{l+k})$$

From this definition we can also define the components of a rank (k,l) tensor T with resepect to a basis $\{v_\mu\}$ for V and $\{v^{\mu *}\}$ for V^* as

$$T = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 * \ldots \nu_l *} v_{\mu_1} \otimes \cdots \otimes v_{\mu_k} \otimes v^{\nu_1 *} \otimes \cdots \otimes v^{\nu_l *}$$

In terms of components, tensor products $S = T \otimes T'$ and contractions have natural definitions

$$C : T^{\ldots\sigma\ldots}{}_{\ldots\sigma\ldots} \qquad S^{\mu_1 \ldots \mu_{k+k'}}{}_{\nu_1 \ldots \nu_{l+l'}} = T^{\mu_1 \ldots \mu_k}{}_{\nu_1 \ldots \nu_l} (T')^{\mu_{k+1} \ldots \mu_{k+k'}}{}_{\nu_{l+1} \ldots \nu_{l+l'}}$$

Now the reason we care about all of this stuff is because of the tangent space V_p at a point p on a manifold M whose elemnts we'll call *contravariant*. We denote V_p^* as the *contagent space* whose elemnts we'll call *covariant*. Now lets find the transformation law for the cotangent space. Recall the defining equation of the dual vector space basis (7)

$$v_\mu v^{\nu *} = \delta_\mu^\nu$$

From this definition and the vector transformation law we can derive the transformation law for the dual basis

$$v_{\mu'} v^{\nu' *} = \frac{\partial x^{\nu'}}{\partial x^{\mu'}} = \frac{\partial x^\nu}{\partial x^\mu}\frac{\partial x^{\nu'}}{\partial x^\nu}\frac{\partial x^\mu}{\partial x^{\mu'}} = v^{\nu *} v_\mu \frac{\partial x^{\nu'}}{\partial x^\nu}\frac{\partial x^\mu}{\partial x^{\mu'}} = v^{\nu *} v_{\mu'} \frac{\partial x^{\nu'}}{\partial x^\nu} \Rightarrow v^{\nu' *} = v^{\nu *}\frac{\partial x^{\nu'}}{\partial x^\nu}$$

Now consider some dual vector $\omega \in V_p^*$. We now have

$$\omega = \omega_{\mu'} v^{\mu' *} = \omega_\mu v^{\mu *} \Rightarrow \boxed{\omega_{\mu'} = \omega_\mu \frac{\partial x^\mu}{\partial x^{\mu'}}}$$

Combining the above result with the vector transformation law naturally leads us to the *tensor transformation law*

$$\boxed{T^{\mu_1' \dots \mu_k'}{}_{\nu_1' \dots \nu_l'} = T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_k} \frac{\partial x^{\nu_1}}{\partial x^{\nu_1'}} \cdots \frac{\partial x^{\mu_1'}}{\partial x_1^{\mu}} \cdots}$$

Lets define a *metric*. The metric is to give the infitesmial displacement squared." Thus it should be a $(0,2)$ tensor, symmetric, and nondegenerate ($g(v, v_1) = 0 \forall v \Leftrightarrow v_1 = 0$). In a coordinate basis

$$g = ds^2 = g_{\mu\nu} dx^\mu \otimes dx^\nu = g_{\mu\nu} dx^\mu dx^\nu$$

We'll prove in the problems that at each point p there exists a orthonormal basis such that $g = \text{diag}(\pm 1, \dots, \pm 1)$. I think what this is saying if we consider space time is that locally any point on a manifold can locally look like flat space.

We can also view the metric as a map from V^p to V^{p*} by only partially evaluating the metric. Since g is nondegenerate the partially applied map $g(v, \dot{)}$ is bijective and thus the metric creates a correspondance between the vector space and the dual vector space.

Abstract Index Notation

Wald uses the conventation that objects with latin indeces (e.g. T^{abc}) are tensors and objects with greek indices (e.g. $T^{\mu\nu\rho}$) are the component representaion of T^{abc} if we introduce a basis. Wald's convention for symmetrization and antisymmetrization are as follows

$$T_{(a_1 \dots a_l)} = \frac{1}{l!}\sum_\pi T_{a_{\pi(1)} \dots a_{\pi(l)}}$$

$$T_{[a_1 \dots a_l]} = \frac{1}{l!}\delta_\pi \sum_\pi T_{a_{\pi(1)} \dots a_{\pi(l)}}$$

Where π is all permutation of $1, \dots, l$ and δ_π is the parity of the permutation π. A *differential form* is a totally antisymmetric tensor field. E.g. $T_{[a_1 \dots a_l]}$ is a l–form.

PROBLEMS

a) We wish to show that $F_{ij}^{\pm\pm} = f_i^{\pm} \circ (f_j^{\pm})^{-1}$. Where the form of f is given

$$f_1^+(x^1, x^2, x^3) = (x^2, x^3) \qquad f_2^+(x_1, x_2, x_3) = (x_1, x_3) \qquad \ldots$$

If $i = j$ we have the identity map or no overlap on the manifold. So we only need to consider $i \neq j$. From the above definitions we can write

$$F_{ij}^{++}(x_i, x_k) = (+x_j, x_k) \qquad F_{ij}^{+-}(x_i, x_k) = (-x_j, x_k) \qquad \ldots$$

Lets consider a simpler case first. For $\mathbb{S}^1$ we can define similar function g and G where

$$G_{ij}^{++}(x_i) = (x_j) \qquad G_{ij}^{-+}(x_i) = (x_j) \qquad \ldots$$

Now consider Figure 2.

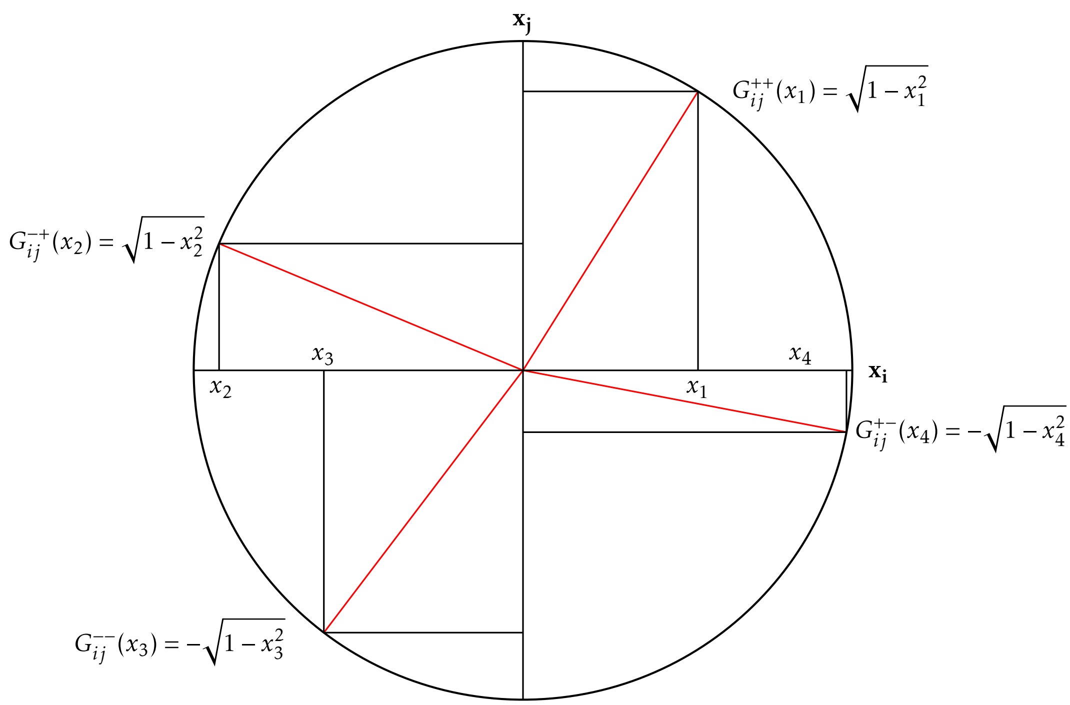


Figure 2: Visualization the function $G_{ij}^{\pm\pm}$ for the $\mathbb{S}^1$ case of Wald Chapter 2 problem 1.

From the figure it should be clear that we're trying to prove that the following function is C^∞.

$$G_{ij}^{\pm}(x_i) = \pm\sqrt{1 - x_i^2}$$

Where the $\pm$ corresponds to the sign of x_j. The equivalent function for $\mathbb{S}^2$ is

$$F_{ij}^{\pm}(x_i, x_k) = \left(\pm\sqrt{1 - x_i^2 - x_k^2}, x_k\right)$$

With the same $\pm$ convention. We only need to consider the first element of the ordered pair because the second element is just an identity map. This function is clearly continous (I don't wanna pull out the $\epsilon - \delta$'s). Also all deriatives would be of the form

$$\frac{d^n}{dx_i^n} F_{ij}^{\pm} = \frac{c_1 x_i^{a_1}}{(1 - x_i^2 - x_k^2)^{\frac{b_1}{2}}} + \frac{c_2 x_i^{a_2}}{(1 - x_i^2 - x_k^2)^{\frac{b_2}{2}}} + \dots$$

And same thing for derivatives with respect to x_k. I know I'm being really hand-wavy here but I think once we get a functional form in terms of x_i and x_k it should be clear that the $f_i^{\pm} \circ (f_j^{\pm})^{-1}$ are C^∞.

b) Now for this problem we can do sort of a wonky construction. Again lets consider the $\mathbb{S}^1$ case. We'll define two coordinate systems f^1 and f^2 which both cover everything except

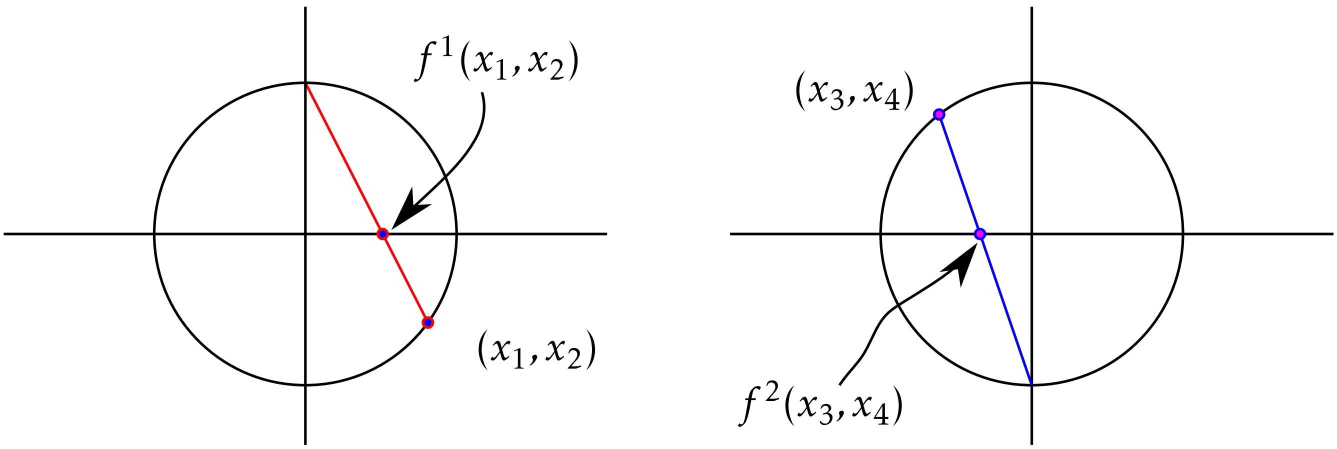


Figure 3: Illustration for two coordinate systems that cover the $\mathbb{S}^1$ which is used in Wald 2.1(b)

for the north and south pole respectiveley. The visualization of these coordinate systems can be found in Figure 3. We now need to find these f^1 and f^2. For f^1 consider drawing a line from $(1, 0)$ to some point on the circle (x_1, x_2) from point slope form we have

$$y = \frac{x_2 - 1}{x_1} x + 1$$

Now we find where this function intersects $y = 0$

$$0 = \frac{x_2 - 1}{x_1} x + 1 \Rightarrow x = -\frac{x_1}{x_2 - 1} = \frac{x_1}{1 - x_2}$$

And we can do something similar for f^2. Clearly the map is bijective as long as we don't include the poles and the overlap is smooth. Now to extend this to the $\mathbb{S}^2$ case we need to find the intersection of a line drawn from the pole to the $x - y$ plane. First lets try to find the x component. This means we need to find the intersection of $z = 0$ and

$$z = 1 + \frac{x_3 - 1}{x_1} x$$

This gives us

$$x = \frac{x_1}{1 - x_3}$$

And similarly we get for the other component

$$y = \frac{x_2}{1 - x_3}$$

So our first map is

$$f^1(x_1, x_3, x_3) = \left\{ \frac{x_1}{1 - x_3}, \frac{x_2}{1 - x_3} \right\}$$

And we don't have to do much to get f^2. Notice that which pole we'r considering is determined by the $1 \pm x_3$ term. For $-$ we consider the north pole so for $+$ we consider the south pole. This means

$$f^2(x_1, x_2, x_3) = \left\{ \frac{x_1}{1 + x_3}, \frac{x_2}{1 + x_3} \right\}$$

The coordinate systems take an open sets of $\mathbb{S}^2$ to $\mathbb{R}^2$ and $f^2(f^1)^{-1}$ is smooth. The bijectiveness of the functions should be verifiable graphically as well as smoothness.

Lecture 2a: Non-Rectangular Coordinates and Tensors

January 25, 2021

Special Relativity in Non-Rectangular Coordinates

Recall the relation between spherical and rectangular coordinates

$$r = (x^2 + y^2 + z^2)^{1/2} \qquad \theta = \arccos(z/r) \qquad \phi = \arctan(y/x)$$

$$\Rightarrow x = r \sin\theta \cos\phi \qquad y = r \sin\theta \sin\phi \qquad z = r \cos\theta$$

To make things simpler we'll consider the 2-d case $x = r\sin\phi$ and $y = r\cos\phi$. Using chain rule we can find how infinitesimal arc lengths work in different coordinate systems

$$dx^2 + dy^2 + \left(\frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \phi} d\phi \right)^2 + \left(\frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \phi} d\phi \right)^2 = dr^2 + r^2 d\phi^2$$

We can extend this to $D = 3 + 1$ where $x^\alpha = x^\alpha(q^\beta)$ and $q^\gamma = q^\gamma(x^\rho)$. We can write the above expression for infinitesimal arc length as

$$\delta_{ij} dx^i dx^j = \delta_{ij} \frac{\partial x^i}{\partial r^b} \frac{\partial x^j}{\partial r^c} dr^b dr^c$$

Lets define the *metric* as

$$g_{bc} = \delta_{ij} \frac{\partial x^i}{\partial r^b} \frac{\partial x^j}{\partial r^c}$$

Straight Lines in 4-D Minkowski Space

A line where someone is only getting older is

$$u^\alpha = (1,0,0,0)$$

And from week one we now another frame moving past this person in the $+x$ direction with speed v the four-velocity of this man is

$$u^{\beta'} = \gamma(1,-v,0,0)$$

The Minkowski metric $\eta_{\mu\nu} = \text{diag}(-1,1,1,1)$ can be used to compute

$$\eta_{\mu'\nu'}u^{\mu'}u^{\nu'} = \gamma^1(-1+v^2) = -1 = \eta_{\mu\nu}u^\mu u^\nu$$

So what we see is that the calculated quantity $\eta_{\mu\nu}u^\mu u^\nu$ is a invariant. We got -1 which is characteristic of a timelike vector. Basically a timelike path in one reference frame is timelike in all reference frames. We can also consider a null vector for example the motion of a photon

$$n^\alpha = (1,1,0,0) \Rightarrow n^{\rho'} = (1+v)^{1/2}(1,1,0,0) \Rightarrow \eta_{\mu\nu}n^\mu n^\nu = 0$$

Lets ask the question: what is a straight timelike line in minkowski space? We'll say it's a path with constant 4-velocity. This means that the path should not be accelerating along any direction which can be stated as

$$u^\alpha u^\beta, \alpha = 0 \tag{2a.1}$$

The comma denotes a partial derivative e.g. $f_{,\rho} = \partial f/\partial x^\rho$. How does this work in non-rectangular coordinate frame? Well we only need to transform the β' coordinates since the α coordinates are dummy indeces (they're summed over)

$$u^{\alpha'}u^{\beta'}{}_{,\alpha'} = \left(\frac{\partial x^{\beta'}}{\partial x^\beta}\right)u^\alpha u^\beta{}_{,\alpha}$$

(Proof for vector transformation law at (6)). An algorithm he gives to calculate this given a non-rectangular coordinate system is as follows

(a) Start with non-rectangular form of tangent vector $u^{\gamma'}$.

(b) Transform to rectangular frame

$$u^\sigma = \frac{\partial x^\sigma}{\partial x^{\lambda'}}u^{\lambda'}$$

(c) Differentiate to get u^σ

$$u^\alpha \frac{\partial}{\partial x^\alpha}u^\sigma = u^\alpha \frac{\partial}{\partial x^\alpha}\left(\frac{\partial x^\sigma}{\partial x^{\lambda'}}u^{\lambda'}\right).$$

(d) Transform back into nonrectangular frame by first noticing

$$u^\alpha \frac{\partial}{\partial x^\alpha} = u^{\gamma'}\frac{\partial}{\partial x^{\gamma'}}.$$

Then the only remaining umprimed index is σ so we transform the whole epxression by $\frac{\partial x^{\rho'}}{\partial x^{\sigma}}$ to bring back to primed frame.

$$\frac{\partial x^{\rho'}}{\partial x^{\sigma}} u^{\gamma'} \left(\frac{\partial x^{\sigma}}{\partial x^{\lambda'}} \frac{\partial u^{\lambda'}}{\partial x^{\gamma'}} + \frac{\partial}{\partial x^{\gamma'}} \left(\frac{\partial x^{\sigma}}{\partial x^{\lambda'}} \right) u^{\lambda'} \right)$$

Now contracting everything together gives us

$$u^{\gamma'} \left\{ \frac{\partial u^{\rho'}}{\partial x^{\gamma'}} + \frac{\partial x^{\rho'}}{\partial x^{\sigma}} \left(\frac{\partial x^{\sigma}}{\partial x^{\gamma'} \partial x^{\lambda'}} \right) u^{\lambda'} \right\} = u^{\gamma'} \left\{ \frac{\partial u^{\rho'}}{\partial x^{\gamma'}} + u^{\lambda'} \Gamma^{\rho'}_{\lambda'\gamma'} \right\} = u^{\gamma'} u^{\rho'}{}_{;\gamma'} = a^{\rho'} = \text{acceleration} \tag{2a.1}$$

This gives us the acceleration and defines a couple of useful things for us. First he semicolon defines the *covariant derivative* ($\nabla_{\lambda'} u^{\rho'}$), and the Γ is the *connection* and is clearly symmetric in its two lower indices by nature of partial derivatives. This also defines parallel transport TODO. The covariant derivative of a scalar is just the partial derivative.

Tensors

Lots of overlap with Wald's Tensors. So lots not typed out here

Contravariant vectors are written with indices up (e.g. $\in V$). The other kind of vector a covariant vector is a covariant vector (e.g. $\in V^*$) which we write with indices lowered. From here we can derive the effect of the covariant derivative on contravariant vectors. Consider a ω_α and y^β and assume that the product rule holes for covariant derivatives.

$$\nabla_\mu(\omega_\alpha y^\alpha) = \omega_\alpha \nabla_\mu y^\alpha + y^\alpha \nabla_\mu \omega_\alpha$$

Now since we know that $\omega_\alpha y^\alpha$ is a scalar we can assert that

$$\omega_\alpha \nabla_\mu y^\alpha + y^\alpha \nabla_\mu \omega_\alpha = \partial_\mu(\omega_\alpha y^\alpha) = \omega_\alpha \partial_\mu y^\alpha + y^\alpha \partial_\mu \omega_\alpha$$

From this we know that the connection contributions must cancel out. This means that

$$\nabla_\mu \omega_\alpha = \omega_{\alpha,\mu} - \omega_\rho \Gamma^\rho_{\mu\alpha}$$

And by extension we can see that for some tensor of rank (k,l) we have that the covariant derivative acts like

$$\nabla_\alpha T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = \left(T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l, \alpha} \right) + T^{\gamma \mu_2 \dots \mu_k}{}_{\nu_1 \dots \nu_l} \Gamma^{\mu_1}_{\gamma\alpha} + \dots - T^{\mu_1 \dots \mu_k}{}_{\gamma \nu_2 \dots \nu_l} \Gamma^{\gamma}_{\nu_1 \alpha} - \dots$$

Now lets try to find a easier way to compute the connection. First consider the covariant derivative on the metric $\nabla_\alpha g_{\mu\nu}$. Going back to our algorithm in the previous section we see that in part (b) we go back to a rectangular frame. This means that the metric becomes the Minkowski metric $\eta_{\mu\nu}$. Then taking the derivative wrt anything is just zero. This means that (2a.1) is equal to zero. Using our form of the covariant derivative's effect on a (k,l) tensor we get

$$0 = g_{\mu\nu,\alpha} - g_{\gamma\nu} \Gamma^\gamma_{\mu\alpha} - g_{\mu\gamma} \Gamma^\gamma_{\nu\alpha}$$

Now lets consider the cyclic permutations

$$0 = g_{\alpha\mu,\nu} - g_{\gamma\mu} \Gamma^\gamma_{\alpha\nu} - g_{\alpha\gamma} \Gamma^\gamma_{\mu\nu}$$

$$0 = g_{\nu\alpha,\mu} - g_{\gamma\alpha}\Gamma^{\gamma}_{\nu\mu} - g_{\nu\gamma}\Gamma^{\gamma}_{\alpha\mu}$$

Using the symmetries of $g_{\mu\nu}$ and the connection we find that when we add the first two and subtracting the last one gives

$$0 = g_{\mu\nu,\alpha} + g_{\alpha\mu,\nu} - g_{\nu\alpha,\mu} - 2g_{\mu\lambda}\Gamma^{\gamma}_{\nu\alpha}$$

Therefore we get the form we desire

$$\Gamma^{\gamma}_{\nu\alpha}g_{\mu\lambda} = \frac{1}{2}\left(g_{\mu\nu,\alpha} + g_{\alpha\mu,\nu} - g_{\nu\alpha,\mu}\right)$$

We can insert at $g^{\mu\lambda}$ on both sides to get

$$\boxed{\Gamma^{\lambda}_{\nu\alpha} = \frac{g^{\mu\lambda}}{2}\left(g_{\mu\nu,\alpha} + g_{\alpha\mu,\nu} - g_{\nu\alpha,\mu}\right)}$$

Lecture 3a: Gravity in Curved Spacetime and The Equivalence Principle

January 26, 2021

Skimming the lecture notes in lecture he covered the equivalence principle, geodesic equation, riemann tensor derivation, Riemann tensor vanish in minkwoski, symmetries of Riemann tensor, I think (13) and (14) are using riemann normal coordinates?, Bianchi identity, Ricci tensor and Ricci scalar, Einstein tensor. Since Matzner's lecture notes are terse we'll have to dig around elsewhere for some sections From Guth's course. The *weak equivalence principle* is the statement that since the force of a gravitational field is equal to $-m\nabla\phi$ where ϕ is the gravitational potential we have that

$$\mathbf{F} = m\mathbf{a} = -m\nabla\phi \Rightarrow \mathbf{a} = -\nabla\phi$$

Namely the acceleration for any object no matter the inertial mass is the same in a gravitational field. We can contrast this with EM where in a electric field the force on a charge q is $F \propto q\nabla V$ where V is the electric potential. The electric charge is not equal to the mass so in an electric field, different particles with different masses and charge will accelerate differently. In this language of charge we could reframe the weak equivalence as the fact that the "gravitational charge" of a object is the same as its mass.

In general relativity we make use a stronger statement called *Einstein's Equivalence Principle* which states that acceleration and local gravitational fields are indistinguishable. If we're in a tiny box we can't tell if we're on the surface of the Earth or on an accelerating space ship. Furthermore since gravitational fields are described by manifolds that look locally flat (e.g. we've associated a tangent space V_p at each point p on our manifold M) we find that locally, experiments will look like they're in flat space (up to $O(s^2)$). %subsectionGeodesic Equation Recall the notion of a "straight line" we defined last lecture (2a.1). We can generalize this for curved spacetime by replacing our regular derivative with a covariant derivative

$$u^{\nu}u^{\mu}{}_{;\nu} = 0$$

This is the directional derivative of some u^μ along the tangent vector u^ν to the curve. More generally any vector w^μ is parallely transported along a curve if

$$v^\mu \nabla_\mu w^\nu = 0$$

Now lets define the Riemann tensor $R^\alpha{}_{\beta\gamma\lambda}$ as

$$\nabla_\gamma \nabla_\lambda u^\alpha - \nabla_\lambda \nabla_\gamma u^\alpha = R^\alpha{}_{\beta\gamma\lambda} u^\beta$$

This Riemann tensor is a precise notion of curvature at a point. To see this consider Figure 4. Let A^μ and B^μ be infinitesmial displacements. If a vector v^μ where transported along A^μ then

$$v^\mu \to v^\mu + A^\nu \nabla_\nu v^\mu$$

Now if we transport v^μ around the entire loop we know that linear terms would vanish since v^μ is parllely transported and all that remains are second order and higher derivative terms which will encode the curvature. Namely

$$\delta v^\mu = A^\alpha B^\beta [\nabla_\alpha, \nabla_\beta] v^\mu$$

From this we can guess that the commutator of the covariant derivatives is what encodes curvature which is why we defined the Riemann tensor as we did. If we explicit write out the

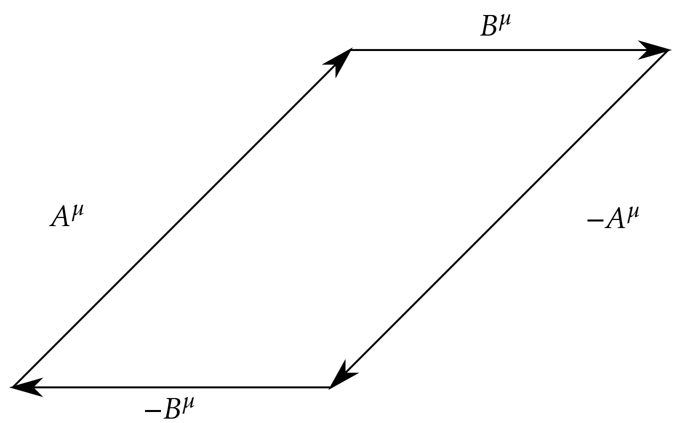


Figure 4: Illustration to motivate the Riemann Curvature tensor

Riemann tensor and grind through some algebra we'll find that

$$R^\alpha{}_{\nu\gamma\lambda} = \Gamma^\alpha{}_{\nu\gamma,\lambda} - \Gamma^\alpha{}_{\nu\lambda,\gamma} + \Gamma^\alpha{}_{\rho\lambda}\Gamma^\rho{}_{\nu\gamma} - \Gamma^\alpha{}_{\rho\gamma}\Gamma^\rho{}_{\nu\lambda}$$

There are several obvious properties of the Riemann curvature tensor.

$$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu}$$

$$R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}$$

$$R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma}$$

And summing cyclic permutations of the last three indices vanish

$$R_{\rho[\sigma\mu\nu]} = 0$$

Also this tensor staisfies the Bianchi idnetity TODO Proof

$$\nabla_{[\mu} R_{\gamma\lambda]\beta\nu} = 0$$

From the Riemann tensor we can also define a few tensors. First the Ricci tensor

$$R_{\nu\lambda} = R^{\gamma}{}_{\nu\gamma\lambda}$$

And the Ricci scalar

$$R = R^{\lambda}{}_{\lambda}$$

Also starting from the Bianchi identity we have

$$\begin{aligned}
\nabla_{[\mu} R_{\gamma\lambda]\beta\nu} &= 0 \\
&= \frac{1}{2}\left(\nabla_\mu R_{\gamma\lambda\beta\nu} - \nabla_\gamma R_{\mu\lambda\beta\nu} - \nabla_\lambda R_{\gamma\mu\beta\nu} + \nabla_\mu R_{\gamma\lambda\beta\nu} - \nabla_\gamma R_{\mu\lambda\beta\nu} - \nabla_\lambda R_{\gamma\mu\beta\nu}\right) \\
&= \nabla_\mu R_{\gamma\lambda\beta\nu} + \nabla_\gamma R_{\lambda\mu\beta\nu} + \nabla_\lambda R_{\mu\gamma\beta\nu}
\end{aligned}$$

Now contracting γ with β leads to

$$\nabla_\mu R_{\lambda\nu} + \nabla^\beta R_{\lambda\mu\beta\nu} - \nabla_\lambda R_{\mu\nu} = 0$$

Now contractin ν with λ gives

$$\nabla_\mu R - \nabla^\beta R_{\mu\beta} - \nabla^\lambda R_{\mu\lambda} = -2\nabla_\alpha G^{\alpha}{}_{\mu} = 0$$

Where we define

$$G^{\alpha}{}_{\mu} = R^{\alpha}{}_{\mu} - \frac{1}{2} R \delta^{\alpha}{}_{\mu}$$

As the *Einstien tensor*

Lecture 4a: Curved Spacetime and Gravity

January 28, 2021

Lets consider small velocities $\frac{v}{c} \ll 1$. In this limit

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$$

The simplest extension of the flat 4-space

$$ds^2 = -(1 + h_{00})dt^2 + dx^2 + dy^2 + dz^2$$

This comes from potentails (e.g. solar system with sun in center.) Lets imagine an asteroid in this gravitaional field. The asteroid folllows geodesics which paths satisfy

$$u^\alpha u^\beta{}_{;\alpha} = u^\alpha(\partial_\alpha u^\beta + \Gamma^\beta_{\sigma\alpha} u^\sigma) = 0$$

First lets consider the green term by considering the christoffel symbol in this limit.

$$\Gamma_{\beta\sigma\alpha} = \frac{1}{2}\left(h_{\beta\alpha,\sigma} + h_{\sigma\alpha,\beta} - h_{\beta\sigma,\alpha}\right)$$

Since h_{00} is the only nonzero element we have that the only nonzero parts of the chtristoffel symbol are

$$\Gamma_{0a0} = \Gamma_{00a} = -\frac{1}{2}h_{00,a} \qquad \Gamma_{a00} = \frac{1}{2}h_{00,a}$$

Now lets find the inverse metric

$$g^{\mu\nu} = \text{diag}\left(\frac{-1}{1+h}1,1,1\right) \approx \text{diag}\left(-1+h,1,1,1\right)$$

We're only taking things first order in h so higher order terms we can ignore. This means for $\Gamma^{\beta}_{\sigma\alpha} = g^{\beta\mu}\Gamma_{\mu\sigma\alpha}$ we can neglect the $\Gamma_{0\ldots}$ terms meaning that the only term we need to care about from the Christoffel symbol comes from $\Gamma^{a}_{00} = \frac{1}{2}\partial^a h_{00}$.

Now consider the red term. In this limit we can neglect higher order terms of v. First consider

$$d\tau^2 = dt^2 - d\mathbf{x}^2 \Rightarrow \left(\frac{d\tau}{dt}\right)^2 = \left(\frac{1}{u^0}\right)^2 = 1 - \frac{d\mathbf{x}^2}{dt^2} = 1 - v^2 \Rightarrow u^0 = \frac{1}{\sqrt{1-v^2}}$$

In our limit we get from a binomial expansion

$$u^0 \approx 1 + \frac{1}{2}v^2 \to \partial_0 u^0 \approx v\frac{dv}{dt}$$

This result means that in our limit we can approximate $u^0 \approx 1$ since v^2 is higher order in v and any term $v^\alpha \partial_\alpha u^0$ we can neglect since they are higher order in v.

$$u^j = \frac{dx}{d\tau} = \frac{dx}{dt}\frac{dt}{d\tau} = \mathbf{v}u^0 = \frac{\mathbf{v}}{\sqrt{1-\mathbf{v}^2}} \approx v + \ldots$$

Thus $u^i\partial_i u^j$ terms also we can neglect since they are higher order in v.

So the only terms in the geodesic equation that are not trivial in this limit are

$$u^0(\partial_0 u^i + \Gamma^i_{00}u^0) \approx \frac{dv^i}{dt} + \frac{1}{2}\partial^i h_{00} = 0 \Rightarrow \boxed{\mathbf{a} = -\frac{1}{2}\nabla h_{00}}$$

If we identify $\frac{1}{2}h_{00}$ with the gravitaional potential then we see this is exactly Newton's equation for motion in a gravitaional field.

I think what Matzner's last section is trying to say is that we want to write a tensor equation that is second derivative in the metric so that it can coincide with with newtonian limit field equation. And after some algebra find that the simplest tensor equation we could write down to satisfy this constraint is

$$\nabla^2 h_{00} = R = 8\pi\rho$$

With $G = c = 1$. This however isn't the Einstein field equation and we'll see what's wrong with the above equation in future lectures.

Lecture 5a: Stress-Energy Tensor and Einstein Equation

February 01, 2021

Lets begin with a definition

Definition 1: (Four-Momentum) Let m be the rest mass of the particle. We define the 4-momentum as

$$p^\mu = mu^\mu = \{\text{Energy}, p^x, p^y, p^z\} = \{m, 0, 0, 0\}$$

This transforms as you would expect under lorentz transformation $p^{\mu'} = \Lambda^{\mu'}{}_\mu p^\mu$. Which gives

$$p^{\mu'} = \{\gamma m, -\gamma m v, 0, 0\}$$

Using this notion of how energy and equivalently mass transforms under loretnz transform we can also consider how mass density transforms under lorentz transformation. Considere n particles each of mass m in some cube V. Once we boost the cube will shrink a factor of gamma γ and the mass (from our definition of four momentum) will grow a factor of γ. This means

$$\rho = \frac{nm}{V} \to \rho' = \gamma^2 \rho$$

But that's weird. If ρ where just a scalar then it would be the same in any frame. What this means is that the equation we wrote down last time

$$R = 8\pi\rho$$

Can't be true since R is a scalar while ρ is something different. This will motivate us to define the stress energy tensor. For now lets consider the simplest case, dust. Just a bunch of particles floating in space

Definition 2: (Stress-Energy Tensor for Dust) Let ρ be the mass density of a cloud of particles *at rest*. The Stress-Energy Tensor $T^{\mu\nu}$ is then defined as

$$T^{\mu\nu} = \rho u^\mu u^\nu$$

Dust is good an all but a good deal of things we care about can't be described by dust. What's missing is pressure. So how do we incorporate pressure into our stress energy tensor? Well pressure arises from molecules crashing into eachother. So maybe if we lorentz boost our Stress-Energy tensor so that the particles are moving we'll get a clue. Consider

$$T^{\mu'\nu'} = T^{\mu\nu}\Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu \Rightarrow T^{1'1'} = \gamma^2 \rho v^2$$

Now lets consider the units of pressure

$$[P] = \left[\frac{\text{Force}}{\text{Area}}\right] = \left[\frac{(\text{Mass}) \times (\text{Accleration})}{(\text{Area})}\right] = \Big[(\text{Density}) \times (\text{Acceleration}) \times (\text{Length}) = [\rho][v^2]\Big]$$

Hmmmmm, intersting, the units of pressure match the units of the boosted spatial component of the Stress-Energy tensor. So perhaps the pressure can be incorporated in the Stress-Energy tensor by putting it as the spatial components. Lets say we have isotropic pressure, then we could guess that in the rest frame

$$T^{\mu\nu} = \text{diag}\{\rho, P, P, P\}$$

How do we write this with tensors so that things are coordinate invariant. Well thinking about it for a little might lead you to

$$T^{\mu\nu} = \rho u^\mu u^\nu + P(\eta^{\mu\nu} + u^\mu u^\nu)$$

Where the green term is how we incoporate the pressure. It's not too hard to show that $\partial_\mu T^{\mu\nu} = 0$. This equation actually encodes both conservation of energy and momentum. So far we've been in flat spacetime but we can bring this to the realm of general relavitity by inserting covariant derivatives and general metrics where needed. The point of all this is to have something we could put on the RHS of a general relativistic field equation that might reduce to the newtonian field equation in the weak gravity limit. Now that we've done this we have

$$G^{\mu\nu} = \kappa T^{\mu\nu}$$

We'll explore what κ is later.

Lecture 5b: Weak Field Gravity

Lets consider the weak gravity limit where the metric is nearly flat for a large portion of ST. In this case

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

For some small perturbation $h_{\mu\nu}$. Clearly under lorentz transformation

$$\eta_{\mu\nu} + h_{\mu\nu} \to \eta_{\mu'\nu'} + h_{\mu'\nu'}$$

Naturally we'd also want to know what the inverse metric is. recall the defining property of the inverse metric

$$g_{\mu\nu} g^{\mu\rho} = \delta^\rho_\nu$$

Lets try to find this inverse metric in the weak field limit. Our first guess might be

$$g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu} \Rightarrow g_{\mu\nu} g^{\mu\rho} = (\eta_{\mu\nu} + h_{\mu\nu})(\eta^{\mu\rho} + h^{\mu\rho}) = \delta^\rho_\nu + \underbrace{h_\nu{}^\rho + h^\rho{}_\nu}_{\text{don't want these guys}} + O(h^2)$$

Well this clearly can't be right. We have some extra linear terms in h that we can't neglect. How do we fix this? Well for starters we could try changing the sign of $h^{\mu\nu}$ so they might cancel. This gives us

$$(\eta_{\mu\nu} + h_{\mu\nu})(\eta^{\mu\rho} - h^{\mu\rho}) = \delta^\rho{}_\nu - h_\nu{}^\rho + h^\rho{}_\nu + O(h^2)$$

But things still aren't quite right. The index ordering is all wrong. So this might inspire us to swap the index placement in $h^{\mu\rho}$ for our inverse metric.

$$(\eta_{\mu\nu} + h_{\mu\nu})(\eta^{\mu\rho} - h^{\rho\mu}) = \delta^\rho{}_\nu - h^\rho{}_\nu + h^\rho{}_\nu + O(h^2) = \delta^\rho{}_\nu + O(h^2)$$

Exactly what we wanted!

Definition 3: (Inverse Metric in Weak Gravity Limit) In the weak gravity limit where only consider first order deviation from flat space time $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ we have the inverse metric

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\nu\mu}$$

We went to all this trouble to get the inverse metric so that we could compute the Ricci Scalar and Tensor which starts with computing the Riemann tensor. The first thing to note is that we can neglect higher order christoffel symbol terms because they are second order in h. Furthermore the derivative of the minkowski metric is zero since $\eta_{\mu\nu}$ is a constant. thus the only terms in our Riemann tensor are the terms which rely on h. I'm not going to write out the computation explicitly (it's just unenlightening algebra after all and we could just use xAct if we wanted to). In the end we get

$$R_{\beta\nu\gamma\lambda} = \frac{1}{2}\left(h_{\beta\lambda,\nu\gamma} + h_{\nu\gamma,\beta\lambda} - h_{\beta\gamma,\nu\lambda} - h_{\nu\lambda,\beta\gamma}\right)$$

Now taking the appropriate trace is decently simple. We apply the inverse metric to our Riemann tensor and ignore the $O(h^2)$ terms leaving us with

$$R^{\beta}{}_{\nu\gamma\lambda} = \eta^{\beta\gamma} R_{\beta\nu\gamma\lambda} = \frac{1}{2}\left(h^{\beta}{}_{\lambda,\nu\beta} + h_{\nu}{}^{\beta}{}_{,\lambda\beta} - h_{,\nu\lambda} - \partial^2 h_{\nu\lambda}\right)$$

Now lets consider infinitesimal displacements of $O(h)$

$$x^\mu \to x^\mu + \xi^\mu$$

The Jacobian we can then compute is

$$\frac{\partial x^\mu}{\partial x^{\nu'}} = \delta^\mu{}_{\nu'} - \xi^\mu{}_{,\nu}$$

Now note that the second term is $O(h)$. If we transform the Riemann tensor to this infinitesimally shifted coordinate system then we'd have the Riemann tensor plus terms $O(h^2)$ which we are neglecting. Thus the Riemann tensor is invariant under infinitesimal coordinate changes in the weak gravity limit. We can also find that

$$g_{\beta'\gamma'} = \eta_{\beta\gamma} + h_{\beta\gamma} - \xi_{\gamma,\beta} - \xi_{\beta,\gamma} \Rightarrow h_{\beta\gamma} \to h_{\beta\gamma} - \xi_{\gamma,\beta} - \xi_{\beta,\gamma}$$

This leads us to define something

Definition 4: (Gauge Transformation in Gravitational Fields) We define a gauge transformation in the weak field limit as

$$h_{\beta\gamma} \to h_{\beta\gamma} - \xi_{\gamma,\beta} - \xi_{\beta,\gamma}$$

This gauge transformation leaves the Riemann tensor invariant (e.g. it is Gauge invariant.) Lets define something else

Definition 5: (Trace Reverse) Given some tensor $T_{\alpha\beta}$ we define the trace reversed tensor as

$$\overline{T}_{\alpha\beta} = T_{\alpha\beta} - \frac{1}{2}T^{\mu}{}_{\mu}\eta_{\alpha\beta}$$

It's clear to see in $D =$ we have $\overline{T} = -T$ where T is the trace of the tensor.

If we write the Ricci tensor and scalar in terms of the trace-reversed $h_{\alpha\beta}$ metric we get

$$R_{\nu\lambda} = \frac{1}{2}\left(\overline{h}^{\beta}{}_{\lambda,\nu\beta} + \overline{h}^{\beta}{}_{\nu,\lambda\beta} + \frac{1}{2}\eta_{\mu\nu}\partial^2\overline{h} - \partial^2\overline{h}_{\nu\lambda}\right) \qquad R = \frac{1}{2}\left(2\overline{h}^{\beta\lambda}{}_{,\lambda\beta} + \partial^2\overline{h}\right)$$

Now recall the gauge freedom we have with the metric deviation h. We'll define a useful gauge

Definition 6: (Lorentz Gauge for Gravitational Waves) We define the Lorentz gauge for gravitational waves as

$$\overline{h}^{\beta}{}_{\alpha,\beta} = 0$$

In the lorentz gauge we can write the Einstein tensor as $G_{\nu\lambda} = -\frac{1}{2}\partial^2\overline{h}_{\nu\lambda}$. This means Einstein's field equation becomes

$$\frac{1}{2}\partial^2\overline{h}_{\nu\lambda} = -\kappa T_{\nu\lambda}$$

It's just a wave equation!

Wald Chapter 3: Curvature

For spaces like $\mathbb{R}^3$ we have an intuitive notion of curvature. If we want to think about the curvature on the surface of a sphere we think about how a vector moves along the surface of a sphere embedded in $\mathbb{R}^3$. This is a notion of *extrinsic curvature*. However the manifolds we're considering are not naturally embedded in any higher dimensional space. So we want to define an *intrinsic notion of curvature*. A notion of curvature that comes from the manifold itself. This intrinsic notion will come from parallel transport. Lets say on the plane we move a vector along a closed path. Once the vector comes back to it's starting position it will be exactly. However if we consider a sphere, this will not be so. Consider Figure 5. But here we run into another problem. On a manifold two distinct points $p, q \in M$ have distinct vector spaces V_p and V_q associated with them. So we need more structure than just what the manifold gives us to define the notion of parallel transport. This extra structure needed (similar to the case of dual spaces and dual-dual spaces) comes from the metric. But more generally what we need is a notion of a derivative. If we know how to take a derivative of a vector field along some path on the manifold then we can learn how to parallel transport a vector along some path. After a bit more thought we can also see that the notion that vectors aren't the same after being parallel transported along some path is the same as saying these derivatives don't commute. We'll introduce all these ideas in this chapter.

Derivative Operators and Parallel Transport

We'll start by defining a covariant derivative.

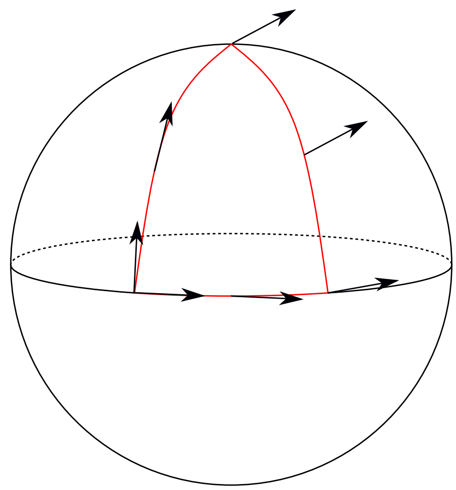


Figure 5: Illustration of parallel transport on the surface of a sphere. Notices that after coming back to its original location, the vector is rotated 90 degrees.

Definition 7: (Covariant Derivative) The covariant derivative $\nabla : \mathcal{T}(k,l) \to \mathcal{T}(k,l+1)$ is a map which satisfies the following properties

(a) Linear

(b) Satisfies a Leibnitz rule

(c) Commutes with contraction (e.g. simples case is $\nabla_d(A^c{}_c) = \nabla_d A^c{}_c$)

(d) Consistent with notion of tangent vectors from scalar fields $t(f) = t^a\partial_a f$ which means that $t(f) = t^a\nabla_a f$ for any $t^a \in V_p$ and function f.

(e) Torsion free: for any function $\nabla_a\nabla_b f = \nabla_b\nabla_a f$.

We usually slap an index directly on the covariant derivative ∇_a but since the covariant derivative isn't a dual vector this is a bit of an abuse of notation.

With this we can define the commutator of two vector v and w. Consider some arbitrary function

$$\begin{aligned}
[v,w](f) &= v(w(f) - w(v)f \\
\text{from (d)} &= v^a\nabla_a(w^b\nabla_b f) - w^a\nabla_a(v^b\nabla_b f) \\
\text{Leibnitz rule} &= v^a\left(w^b\nabla_a\nabla_b f + (\nabla_b f)(\nabla_a w^b)\right) - w^a\left(v^b\nabla_a\nabla_b f + (\nabla_b f)(\nabla_a v^b)\right) \\
\text{Use (e) to cancel red terms} &= \left(v^a(\nabla_a w^b) - w^a(\nabla_a v^b)\right)(\nabla_b f)
\end{aligned}$$

From this we can read off

$$\boxed{[v,w]^b = (v^a\nabla_a w^b - w^a\nabla_a v^b)}$$

Like most things in life, we now need to prove existence and uniqueness of the derivative operator. Existence is easy. Just think of the regular partial derivative operator. However this operator is coordinate system dependent(e.g. we define the partial derivative operator in terms of some coordinate system ψ) and isn't associated with the structure of our manifold.TODO: Proof of the statment "given any two derivative operators $\tilde{\nabla}_a$ and ∇_a there exists a tensor field $C^c{}_{ab}$ where

$$\nabla_a \omega_b = \tilde{\nabla}_a \omega_b - C^c{}_{ab}\omega_c$$

Lets explore what this $C^c{}_{ab}$ means for us. First we can show that C is symmetric in its lower two indices. In the above equation let $\omega_b = \nabla_b f$. Since the two different derivative operators must agree on their action on scalar fields form (d) of the definition, we have $\omega_b = \nabla_b f = \tilde{\nabla}_b f$. This gives us

$$\nabla_a \nabla_b f = \tilde{\nabla}_a \tilde{\nabla}_b f - C^c{}_{ab}\nabla_c f$$

Now since the red terms are symmetric under interchange $a \leftrightarrow b$ we know that $C^c{}_{ab}$ must also be symmetric under interchange of two indices. From here we can now show that the C tensor field determines the deviation of the activation between two derivative operators for any arbitrary tensor. Lets start with the next non-trivial case. Consider $\omega_b t^b$. We know that since this is a scalar we have

$$(\nabla_a \omega_b t^b - \tilde{\nabla}_a \omega_b t^b) = 0$$

We can also use the chain rule to get that

$$(\nabla_a - \tilde{\nabla}_a)(\omega_b t^b) = t^b(\nabla_a - \tilde{\nabla}_a)\omega_b + \omega_b(\nabla_a - \tilde{\nabla}_a)t^b$$

We already know what happesn to the first term from the pink term above. So now all we have is

$$-t^b C^c{}_{ab}\omega_c + \omega_b(\nabla_a - \tilde{\nabla}_a)t^b = 0$$

On the first term we can relabel dummy indces $b \leftrightarrow c$ to get

$$\omega_b((\nabla_a - \tilde{\nabla}_a)t^b - t^c C^b{}_{ac}) = 0 \Rightarrow \underline{\nabla_a t^b = \tilde{\nabla}_a t^b + C^b{}_{ac}t^c}$$

And we can keep doing simlar procedures again and again to show that

$$\nabla_a T^{b_1 \ldots b_k}{}_{c_1 \ldots c_l} = \tilde{\nabla}_a T^{b_1 \ldots b_k}{}_{c_1 \ldots c_l} + C^{b_1}{}_{ac} T^{c \ldots b_k}{}_{c_1 \ldots c_l} + \cdots - C^c{}_{c_1 b} T^{b_1 \ldots b_k}{}_{c \ldots c_l} - \ldots$$

All this song and dance is to show that the difference between two derivative operators is characterized by this $C^c{}_{ab}$ tensor we've found. In the special case when one of the derivative operators we're considering if the partial derivative that's associated with a coordinate system then $C^c{}_{ab}$ becomes the *christoffel connection*.

$$\nabla_a \omega_b = \partial_a \omega_b - \Gamma^c{}_{ab}\omega_c$$

Something to note about this christoffel symbol is that it's determined by the derivative operator ∂_a and which is associated with a coordinate system. If we change coordinate systems, the partial derivative changes. When the partial derivative changes we get a completley new Christoffel symbol. Furthermore the components of the christoffel symbol are written in a different coordinate system. Since we've changed coordinates as well as the object of a Christoffel symbol, the christoffel symbol doesn't transform under the tensor transformation law[1].

[1] I don't think this explanation reads very well. TODO

We can now define the notion of parallel transport. A vector v^a is parallely transported around a curve C with tangent vector t^a if along the curve the following equation is satisfied

$$t^a\nabla_a v^b = 0$$

What this equation is saying is that the directional derivative of v^b along the tangent vector to the curve t^a is zero. If we choose a coordinate system we can rewrite the above in terms of our Christoffel connection and regular partial derivative

$$t^a\nabla_a v^b = t^a\partial_a v^b + t^a\Gamma^b{}_{ac}v^c = 0$$

From here we can parameterize the curve with a parameter t which will give us

$$\frac{dv^\nu}{dt} + t^\mu\Gamma^\nu{}_{\mu\rho}v^\rho = 0$$

The above is a differential equation with a unique solution if the inital condition is specified. What this means is that along the curve we have a unique notion of a parallely transported vector. This is important because it's saying if we have a derivative operator and a curve connecting two points on a manifold p and q then we can define a map from V_p to V_q which will just be the vector in V_p parallely transported to V_q which is governed by the differential equation above. But we still have a problem, we have a lot of different derivative operators to choose from. To resolve this we'll show in just a second that we can use the metric to define a unique derivative operator. Specifically we'll require that the inner product remains unchanged as we parellely transport some vectors along any curve.

$$t^a\nabla_a(g_{bc}v^b w^c) = 0$$

By leibnitz rule we have

$$g_{bc}v^b\underbrace{t^a\nabla_a w^c}_{0} + g^{bc}w^c\underbrace{t^a\nabla_a v^b}_{0} + v^b w^c t^a\nabla_a g_{bc} = 0 \Rightarrow \nabla_a g_{bc} = 0$$

We call this property metric compatibility. To show that this defines a unique covariant derivative we'll explicitly show that this condition implies a unique $C^a{}_{bc}$. Consider two possible derivative operators ∇_a and $\tilde{\nabla}_a$. By nature of derivative operators we have

$$\nabla_a g_{bc} = \tilde{\nabla}_a g_{bc} - C^d{}_{ab}g_{dc} - C^d{}_{ac}g_{bd}$$ [2]

The blue term vanishes by metric compatibility leaving us with

$$\tilde{\nabla}_a g_{bc} = C^d{}_{ab}g_{dc} + C^d{}_{ac}g_{bd}$$

By summing and subtracting cyclic permutations of the above and using the symmetry of $C^a{}_{bc}$ we can get(I think we did this in one of the lectures so I won't write it out explicitly)

$$C^c{}_{ab} = \frac{1}{2}g^{cd}\left\{\tilde{\nabla}_a g_{bd} + \tilde{\nabla}_b g_{ad} - \tilde{\nabla}_d g_{ab}\right\}$$

Thus the metric connection uniquely determines $C^c{}_{ab}$. In particular we'll choose $\tilde{\nabla}_a$ to be the partial derivative operator whic means that

$$\Gamma^\lambda{}_{\mu\nu} = \frac{1}{2}g^{\lambda\sigma}\left\{\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right\}$$

[2] hehe

Curvature

We showed that the connection between a vector $p \in V_p$ and $q \in V_q$ can be defined by parallely transporting that vector along some curve. However the $q \in V_q$ will be determined by what curve we pick to transport p along. This will be our notion of curvature which we'll encapsulate in the Riemann curvature tensor that will be constructed in this section.

First lets do some algebra again. Consider two derivative operators ∇_a and ∇_b as well as a function f and dual vector $\omega_c \in V_p^*$. Consider

$$\begin{aligned}(\nabla_a\nabla_b)(f\omega_c) &= \nabla_a(f\nabla_b\omega_c + \omega_c\nabla_b f) \\ &= (\nabla_b\omega_c)(\nabla_a f) + f\nabla_a\nabla_b\omega_c + (\nabla_a\omega_c)(\nabla_b f) + \omega_c\nabla_a\nabla_b f\end{aligned}$$

If we subtract $\nabla_b\nabla_a(f\omega_c)$ from the above then the red terms will cancel leaving us with

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)(f\omega_c) = f\nabla_a\nabla_b\omega_c$$

So from this we can see that $(\nabla_a\nabla_b - \nabla_b\nabla_a)$ is a operator that takes in a $(0,1)$ tensor and outputs a rank $(0,3)$ tensor. From this we can see that this makes $(\nabla_a\nabla_b - \nabla_b\nabla_a)$ a $(1,3)$ tensor meaning that we can define ${R_{abc}}^d$ as

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c = {R_{abc}}^d\omega_d$$

We call this the *Riemann curvature tensor*

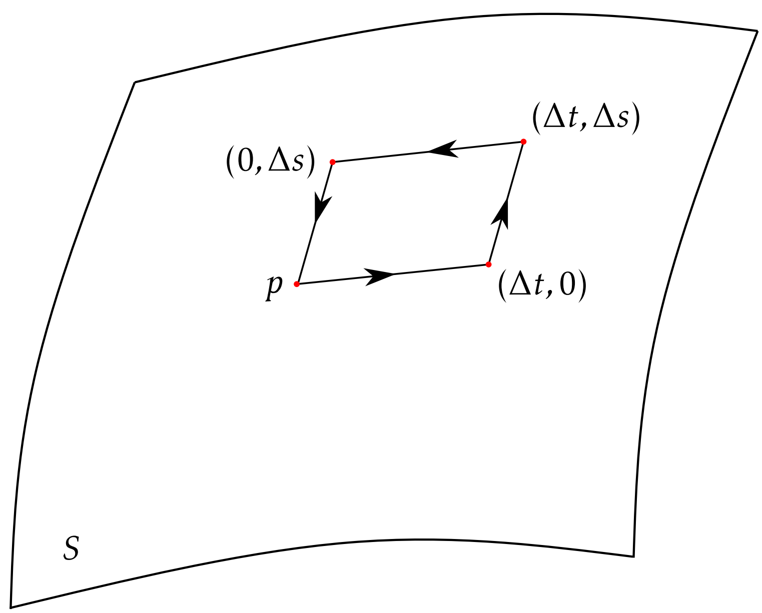


Figure 6: The path we'll parallel transport of some vector v^a on a surface S with coordinates (s,t) on the manifold around some point p. We'll use this to prove some things about the Riemann curvature tensor

Now lets show that this Riemann curvature tensor corresponds to the failure ofa vector to remain invariant under parallel transport. Consider the path in Figure 6 a vector could be parallely transported on. We're considering a surface S with coordinates (s,t)(with $p = (0,0)$)

on the manifold M around some point $p \in M$. It'll be easier to find the change in some vector v^a after being parallely transported around the loop if we also consider a arbitrary dual vector field ω_a and find the change of $v^a\omega_a$. So for the first leg of the journey when we go from p to $(\Delta t, 0)$ we have

$$\delta_1 = \Delta t \partial_t (v^a \omega_a)\Big|_{\left(0, \frac{\Delta t}{2}\right)}$$

We evaluate the derivative at the middle of the transport so that we're only off by factors $O(\Delta t^2)$.[3] We can rewrite ∂_t as $T^a\nabla_a$, the directional derivative of $v^a\omega_a$ along the tangent vector T along the curve. Notice we can also use the defining equation for a parallely transported vector $t^a\nabla_a v^b = 0$ to simplify thigns

$$\delta_1 = \Delta t T^b \nabla_b (v^a\omega_a)\Big|_{(0,\Delta t/2)} = \Delta t v^a T^b \nabla_b \omega_a\Big|_{(0,\Delta t/2)}$$

By a similar logic we can also get

$$\delta_2 = \Delta s v^a S^b \nabla_b \omega_a\Big|_{(\Delta s/2, \Delta t)} \qquad \delta_3 = -\Delta t v^a T^b \nabla_b \omega_a\Big|_{(\Delta s, \Delta t/2)} \qquad \delta_4 = -\Delta s v^a S^b \nabla_b \omega_a\Big|_{(\Delta s/2, 0)}$$

Now lets consider the quantity

$$\delta_1 + \delta_3 = \Delta t \left(v^a T^b \nabla_b \omega_a\Big|_{(0,\Delta t/2)} - v^a T^b \nabla_b \omega_a\Big|_{(\Delta s,\Delta t/2)} \right)$$

Now to first order as $\Delta s \to 0$ we find that the above quantity vanishes and something similar happens for $\delta_2 + \delta_4$ so to first order there is no change in $v^a\omega_a$. So we need to consider higher order effects. Our next order correction is explicitly parallely transporting v^a and $T^b\nabla_b\omega_a$ up Δs by asserting

$$v^a\Big|_{\Delta s, \Delta t/2} = v^a\Big|_{0,\Delta t/2} + \Delta s S^c \nabla_c v^a\Big|_{\Delta s/2, \Delta t/2}$$

And a simlar expresion for $T^b\nabla_b\omega_a$. However note that to first order the above term vanishes and only the $T^b\nabla_b\omega_a$ term changes thus we have(while ignoring the big ugly "evaluate at" symbols since they don't really mean anything right now)

$$\delta_1 + \delta_3 = -\Delta t \Delta s v^a S^c \nabla_c T^b \nabla_b \omega_a$$

By a similar logic we can also get (notice the sign difference)

$$\delta_2 + \delta_4 = \Delta t \Delta s v^a T^c \nabla_c S^b \nabla_b \omega_a$$

[3] accordint to Wald. I haven't confirmed this myself. I'll take Wald's word here sense proving it seems extraneous to the proof right now.

Thus in total we can find that

$$\begin{aligned}
\delta(v^a\omega_a) &= \delta_1 + \delta_3 + \delta_2 + \delta_4 \\
&= \Delta t \Delta s v^a \left(T^c \nabla_c S^b \nabla_b \omega_a - S^c \nabla_c T^b \nabla_b \omega_a \right) \\
&= \Delta t \Delta s v^a \left(T^c \left(S^b \nabla_c (\nabla_b \omega_a) + (\nabla_b \omega_a)(\underbrace{\nabla_c S^b}_{\substack{=0 \text{ because} \\ S^b \text{ is constant}}}) \right) - S^c \left(T^b \nabla_c (\nabla_b \omega_a) + (\nabla_b \omega_a)(\underbrace{\nabla_c T^b}_{=0}) \right) \right) \\
&= \Delta t \Delta s v^a T^c S^b \left(\nabla_c \nabla_b - \nabla_b \nabla_c \right) \omega_a \\
&= \Delta t \Delta s v^a T^c S^b R_{cba}{}^d \omega_d
\end{aligned}$$

From this we can see that the Riemann curvature tensor is exactly the failure for a vector to remain invariant under parallel transport on some closed loop. To see this more explicitly consider

$$\delta(v^a\omega_a) = \delta(v^a)\omega + v^a \underbrace{\delta(\omega_a)}_{=0}$$

This expicitly gives us

$$\delta(v^a) = \Delta t \Delta s v^d T^c S^b R_{cbd}{}^a$$

TODO 3.2.10,3.2.11,3.2.12