

Contents

0.1	Introduction	
0.1.1	Introduction to General Relativity	
0.1.2	Introduction to Compact Objects and Boson Stars	
0.1.3	Conventions	
0.2	Differential Geometry	
0.2.1	Introduction to Geometry and Manifolds	
0.2.2	Functions, Curves and Tensors on Manifolds	
0.2.3	The Inner Product and the Metric	
0.2.4	Maps Between Manifolds	
0.2.5	Lie Derivatives	
0.2.6	Lengths on Manifolds	
0.2.7	Volumes on Manifolds	
0.2.8	Geodesics	
0.3	Tensor Calculus and Curvature	
0.3.1	General Covariance and Coordinate transformations	
0.3.2	The Covariant Derivative	
0.3.3	The Connection	
0.3.4	Curvature Tensors	
0.3.5	The Divergence Theorem	
0.4	Relativity	
0.4.1	Special Relativity	
0.4.2	Physics in Special Relativity	
0.4.3	Physics in Curved Space	
0.4.4	The Stress-Energy-Momentum Tensor	
0.4.5	The Einstein Equation	
0.4.6	Black Holes	
0.4.7	The Cosmological Constant	
0.4.8	The Lagrangean Formulation of General Relativity	
0.5	Stuff	
0.5.1	notes	
0.5.2	Differential Forms	
0.5.3	Parallel Transport	
0.5.4	Low Curvature Limit of General Relativity	

0.1 Introduction

0.1.1 Introduction to General Relativity

general gr shit? mention einstein derivation (hilbert derivation a potentially earlier but vacuum), karl schwarzschild (ironically meaning black shield), then low mass limit, photon deflection and mercury perihelion around sun. mention cosmology, gravitational waves, more black holes, minkowski. more compact objects.

0.1.2 Introduction to Compact Objects and Boson Stars

The first non-trivial solution to Einstein's equation found was that of the spherically symmetric, static and asymptotically flat vacuum spacetime by Karl Schwarzschild in 1915. The solution was designed to be used outside a spherically symmetric, non-spinning, body of mass; however it turned out to provide use in describing black holes. This metric was then modified by Tolman, Oppenheimer and Volkov in 1939 to describe the non-vacuum case of a constant density neutron star. This turned out to give an unphysical estimate of $0.7M_{\odot}$ for the upper limit of neutron star mass due to the equation of state.

The study of compact exotic objects can be traced back to John Wheeler who investigated Geons in 1955 for their potential similarity to elementary particles. Geons are gravito-electromagnetic objects with the name arising from "gravitational electromagnetic entity". In 1968 David Kaup published [1] describing what he called "Klein-Gordon Geons", nowadays referred to as boson stars. Importantly, boson stars are a localised complex Klein-Gordon configuration, with the real counterparts being unstable. Many variants such as (Spin 1) proca stars [2], electromagnetically charged boson stars and many others have been studied.

Interest in boson stars remains for many reasons. Given the recent discovery of the higgs boson, we know that scalar fields exist in nature and any gravitational wave signals created by compact objects could theoretically be detected with modern gravitational wave interferometers. Secondly, boson stars are a good candidate for dark matter haloes. Boson stars are also useful as a proxy to other compact objects in general relativity; there is a lot of freedom in the construction of different types of boson star and they can be fine tuned to model dense neutron stars for one example. The advantage this would have over simulating a real fluid is that the Klein Gordon equation is linear in the principal part meaning smooth data must always remain smooth; thus avoiding shocks and conserving particle numbers relatively well with less sophisticated numerical schemes.

On a slightly different topic, collisions of boson stars could be a natural method to produce scalar hair around black holes which will be discussed later in more detail.

0.1.3 Conventions

Throughout this thesis physical quantities will be expressed as a dimensionless ratio of the Planck length, time and mass L_{pl} , T_{pl} and M_{pl} respectively; consequently the constants c , G and \hbar evaluate numerically to 1. As an example, Newtons equation of gravity would be recast like

$$F = \frac{GMm}{r^2} \rightarrow \left(\frac{F}{F_{pl}} \right) = \frac{\left(\frac{M}{M_{pl}} \right) \left(\frac{m}{M_{pl}} \right)}{\left(\frac{r}{L_{pl}} \right)^2} \quad (0.1.1)$$

where $F_{pl} = M_{pl}L_{pl}T_{pl}^{-2}$ is the Planck force. $c = G = \hbar = 1$, unless stated otherwise. The metric signature will always be $(-, +, +, +)$.

Tensors and tensor fields will be denoted using bold font for index free notation and normal font for the components. The dot product between two vectors or vector fields will be written interchangeably as $\mathbf{A} \cdot \mathbf{B} \leftrightarrow A_{\mu}B^{\mu}$ for readability. Additionally, ∇_{μ} denotes the covariant derivative and ∂_{μ} is the partial derivative, both with respect to coordinate x^{μ} .

When considering the ADM decomposition, as in [REF SECT], objects can be associated with both the 3+1 dimensional manifold \mathcal{M} or the 3 dimensional hypersurface Σ . To differentiate here, standard Roman letters such as R represent the object belonging to \mathcal{M} and calligraphic letters such as \mathcal{R} correspond to the projected object belonging to Σ . [MAYBE JUST REMOVE THIS BIT AND MAKE IT OBVIOUS IN THE ACTUAL 3+1 SECTION].

Finally, unless stated otherwise, Greek indices such as $\{\alpha, \beta, \dots, \mu, \nu, \dots\}$ label four dimensional tensor components whereas late Latin indices such as $\{i, j, k, \dots\}$ label three dimensional tensor components and early Latin indices such as $\{a, b, \dots\}$ label two dimensional ones. When the index range is unspecified and unimportant Greek letters will also be used.

EINSTEIN SUMMATION CONV

make explicit th inner product, dot product, outer product (otimes) and wedge product (for forms, antisymm)

0.2 Differential Geometry

0.2.1 Introduction to Geometry and Manifolds

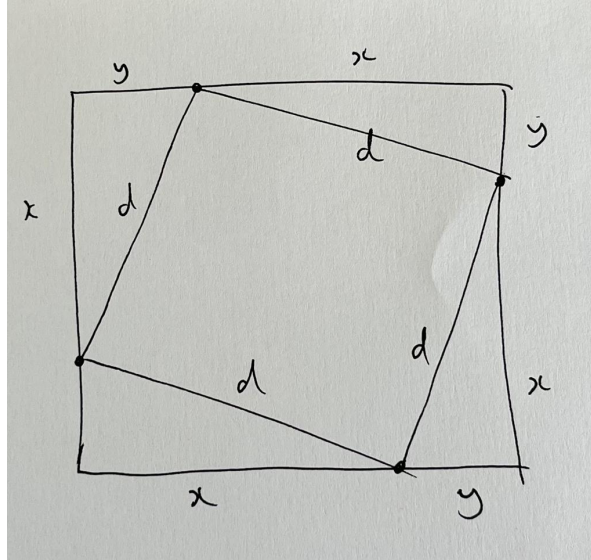


Figure 1: Diagram for proof of Pythagoras' theorem.

Everyones first encounter with geometry will cover Pythagoras' theorem; arguably the most famous and useful equation in existance. Pythagoras' equation relates the sidelengths of a right angled triangle, it says that $s^2 = x^2 + y^2$ for a triangle with height y , width x and hypotenuse length s . This can be shown very simply by looking at Fig. 1. The area of the partially rotated square is s^2 , but we can also calculate it from the the area of the larger square A_{sq} and subtracting four times the area of one of the triangles A_{tr} . Given that $A_{sq} = (x + y)^2$ and $A_{tr} = \frac{1}{2}xy$, then

$$s^2 = (x + y)^2 - 2xy = x^2 + y^2, \quad (0.2.1)$$

and we have proved Pythagoras' theorem. Using an infinitessimally small triangle, we can write $ds^2 = dx^2 + dy^2$ and this can be trivially extended to arbitrary dimensions like

$$ds^2 = dx^2 + dy^2 + dz^2 + \dots \quad (0.2.2)$$

The infinitesimal form of Pythagoras' theorem is very powerful as it can be used to calculate the length of a generic curve by approximating the curve as a collection of infinitessimally small straight lines with

length ds . So far we have assumed that space is flat meaning Eq. (0.2.2) is true for all points in space, this is an assumption we will have to drop if we want to study the curved spaces arising in strong gravity. In the next sections we will explore the generalisation of Pythagoras' equation to curved spaces and use it to measure curve lengths as well as volumes and areas.

Differential Geometry (DG) is the extension of calculus, linear algebra and multilinear algebra to general geometries. Einstein's Theory of Relativity is written using the language of DG as it is the natural way to deal with curves, tensor calculus and differential tensor equations in curved spaces. For a basic introduction to DG, we should start with a manifold \mathcal{M} which is an N dimensional space that locally looks like \mathbb{R}^N , N dimensional Euclidean space. This is important as at a point $p \in \mathcal{M}$ we can find infinitesimally close neighbouring points $p + \delta p \in \mathcal{M}$. In the following sections we will explore curves, functions, tensors and calculus on manifolds using DG.

0.2.2 Functions, Curves and Tensors on Manifolds

A real scalar function f over \mathcal{M} maps any point $p \in \mathcal{M}$ to a real number, this is denoted as $f : p \rightarrow \mathbb{R}$. An important example of a set of scalar functions is the coordinate system ϕ , $\phi : p \rightarrow \mathbb{R}^N$, this is normally written x^μ where $\mu \in \{0, 1, \dots, N-1\}$ is an index labelling the coordinate. The map ϕ is called a chart, and unlike Euclidean space one chart may not be enough to cover the entire manifold; in this case a set of compatible charts should be smoothly joined, collectively known as an atlas.

Now that functions have been discussed, the next simplest object we can discuss is a curve, or path, through \mathcal{M} . A curve Γ is a set of smoothly connected points $p(\lambda) \in \mathcal{M}$ that smoothly depend on an input parameter $\lambda \in [\lambda_0, \lambda_1]$. This can be expressed in terms of coordinates as $x^\mu(\lambda)$ where $\phi : p(\lambda) \rightarrow x^\mu(\lambda)$. Differentiating a function f along Γ with respect to λ gives

$$\frac{d}{d\lambda} f(x^\mu(\lambda)) = \frac{dx^\nu}{d\lambda} \frac{\partial f(x^\mu)}{\partial x^\nu} = \frac{dx^\nu}{d\lambda} \partial_\nu f, \quad (0.2.3)$$

where the Einstein summation convention was invoked, summing over all values of ν , and $\partial_\nu = \partial/\partial x^\nu$. Equation (0.2.3) was derived independantly of the choice of f , therefore we can generally write

$$\frac{d}{d\lambda} = \frac{dx^\nu}{d\lambda} \partial_\nu. \quad (0.2.4)$$

The operator $d/d\lambda$ can act on any function f and return a new function \tilde{f} over \mathcal{M} , this is written as $d/d\lambda(f) = \tilde{f}$ where $\tilde{f} : p \rightarrow \mathbb{R}$ for $p \in \mathcal{M}$. We can also think of $d/d\lambda$ as a vector \mathbf{X} with components $X^\mu = dx^\mu/d\lambda$ and basis vectors $\mathbf{e}_\mu := \partial_\mu$ taken from Eq. (0.2.4). The vector \mathbf{X} can be written as $\mathbf{X} = X^\mu \mathbf{e}_\mu$ and can act on a general function f over \mathcal{M} as $\mathbf{X}(f) = X^\mu \mathbf{e}_\mu(f) = X^\mu \partial_\mu f$. Considering the set of all possible curves through a points $p \in \mathcal{M}$, the tangent vector components $dx^\mu/d\lambda$ span an N dimensional space with basis $\mathbf{e}_\mu = \partial_\mu$; this space is called the tangent space and is denoted as $\mathcal{T}_p(\mathcal{M})$ at a point $p \in \mathcal{M}$.

The next object to discuss is the co-vector which is defined as a map from vectors to real numbers; not to be confused with the dot product in section 0.2.3. Similarly to vectors, a co-vector $\boldsymbol{\omega}$ can be expressed as a sum of components ω_μ and basis co-vectors $\boldsymbol{\theta}^\mu$ like $\boldsymbol{\omega} = \omega_\mu \boldsymbol{\theta}^\mu$. Contrary to vectors, co-vector components have downstairs indices and the basis has upstairs indices; this choice improves the readability of tensor equations when working with components. A co-vector can map a vector to a real number like $\boldsymbol{\omega} : \mathbf{X} \rightarrow \mathbb{R}$ or $\boldsymbol{\omega}(\mathbf{X}) \rightarrow \mathbb{R}$. Vectors are equally able to map co-vectors to real numbers, denoted as $\mathbf{X} : \boldsymbol{\omega} \rightarrow \mathbb{R}$. Co-vectors are defined such that $\boldsymbol{\theta}^\mu : \mathbf{e}_\nu = \delta^\mu_\nu$ where δ^μ_ν are the components of the Kroneka delta equating to zero unless $\mu = \nu$ in which case they equal unity. The operation of a generic co-vector $\boldsymbol{\omega}$ on a generic vector \mathbf{X} is

$$\boldsymbol{\omega} : \mathbf{X} = \omega_\mu X^\nu \boldsymbol{\theta}^\mu : \mathbf{e}_\nu = \omega_\mu X^\nu \delta^\mu_\nu = \omega_\mu X^\mu \in \mathbb{R}. \quad (0.2.5)$$

This map is linear and identical under reversing the order of operation; $\boldsymbol{\omega} : \mathbf{X} = \mathbf{X} : \boldsymbol{\omega}$. Similarly to vectors, the set of all possible co-vectors at a point $p \in \mathcal{M}$ span an N -dimensional space called the co-tangent space, written as $\mathcal{T}_p^*(\mathcal{M})$.

Multilinear Maps and Tensors

Generalising the previous linear maps between vectors and co-vectors gives the multilinear map. Consider a tensor \mathbf{T} , this can be expressed in component form like

$$\mathbf{T} = T_{\mu\nu,\dots}^{\alpha\beta,\dots} \mathbf{e}_\alpha \otimes \mathbf{e}_\beta \otimes \dots \otimes \boldsymbol{\theta}^\mu \otimes \boldsymbol{\theta}^\nu \otimes \dots \quad (0.2.6)$$

for an arbitrary number of outer products of vector and co-vector bases. A tensor with m co-vector bases and n vector bases is called an (m, n) tensor and has a rank of $m + n$. Vectors, co-vectors and scalars are $(1, 0)$, $(0, 1)$ and $(0, 0)$ tensors respectively. Tensors can act as multilinear maps between tensors. We have already seen how a vector and co-vector can map each other to a scalar, let's extend this with an example. An $(0, 2)$ tensor, $\mathbf{T} = T_{\mu\nu} \boldsymbol{\theta}^\mu \otimes \boldsymbol{\theta}^\nu$ at $p \in \mathcal{M}$, can map two vectors \mathbf{X} and \mathbf{Y} to a scalar as shown,

$$\mathbf{T}(\mathbf{X}, \mathbf{Y}) = T_{\mu\nu} X^\alpha Y^\beta \boldsymbol{\theta}^\mu \otimes \boldsymbol{\theta}^\nu (\mathbf{e}_\alpha, \mathbf{e}_\beta), \quad (0.2.7)$$

$$= T_{\mu\nu} X^\alpha Y^\beta (\boldsymbol{\theta}^\mu : \mathbf{e}_\alpha) (\boldsymbol{\theta}^\nu : \mathbf{e}_\beta), \quad (0.2.8)$$

$$= T_{\mu\nu} X^\alpha Y^\beta \delta_\alpha^\mu \delta_\beta^\nu, \quad (0.2.9)$$

$$= T_{\mu\nu} X^\mu Y^\nu. \quad (0.2.10)$$

The multilinear map can also output generic tensors, for example consider

$$\mathbf{T}(\mathbf{X}, \star) = T_{\mu\nu} X^\alpha (\boldsymbol{\theta}^\mu : \mathbf{e}_\alpha) \boldsymbol{\theta}^\nu = T_{\mu\nu} X^\mu \boldsymbol{\theta}^\nu, \quad (0.2.11)$$

which uses the $(0, 2)$ tensor \mathbf{T} to map the vector \mathbf{X} to a co-vector \mathbf{W} with components $W_\mu = T_{\mu\nu} X^\nu$.

One final example of a mapping is from a single tensor to a lower rank tensor, this is called contraction. To illustrate this, let's take a $(1, 3)$ tensor $\mathbf{Z} = Z_{\mu\nu\rho}^\alpha \mathbf{e}_\alpha \otimes \boldsymbol{\theta}^\mu \otimes \boldsymbol{\theta}^\nu \otimes \boldsymbol{\theta}^\rho$. We can choose to use the basis vector \mathbf{e}_α to act on any of the three co-vector bases, choosing $\boldsymbol{\theta}^\mu$ this is

$$Z_{\mu\nu\rho}^\alpha (\mathbf{e}_\alpha : \boldsymbol{\theta}^\mu) \boldsymbol{\theta}^\nu \otimes \boldsymbol{\theta}^\rho = Z_{\mu\nu\rho}^\mu \boldsymbol{\theta}^\nu \otimes \boldsymbol{\theta}^\rho = \tilde{Z}_{\nu\rho} \boldsymbol{\theta}^\nu \otimes \boldsymbol{\theta}^\rho \quad (0.2.12)$$

where $\tilde{Z}_{\nu\rho} = Z_{\mu\nu\rho}^\mu$.

0.2.3 The Inner Product and the Metric

To introduce the notion of length on a tangent plane $\mathcal{T}_p(\mathcal{M})$ at a point $p \in \mathcal{M}$ the metric tensor \mathbf{g} is introduced. The metric tensor is

$$g_{\mu\nu} = \mathbf{e}_\mu \cdot \mathbf{e}_\nu \quad (0.2.13)$$

where $\mathbf{e}_\mu \cdot \mathbf{e}_\nu$ represents the inner product (or dot product) on $\mathcal{T}_p(\mathcal{M})$; clearly the metric is symmetric by construction as $\mathbf{e}_\mu \cdot \mathbf{e}_\nu = \mathbf{e}_\nu \cdot \mathbf{e}_\mu$. The inner product can be thought of as a multilinear map,

$$\mathbf{g} : (\mathbf{X}, \mathbf{Y}) \rightarrow \mathbb{R}, \quad (0.2.14)$$

$$\mathbf{g}(\mathbf{X}, \mathbf{Y}) = g_{\mu\nu} X^\mu Y^\nu, \quad (0.2.15)$$

where $\mathbf{X} \in \mathcal{T}_p(\mathcal{M})$, $\mathbf{Y} \in \mathcal{T}_p(\mathcal{M})$ and $\mathbf{g} \in \mathcal{T}_p^*(\mathcal{M}) \otimes \mathcal{T}_p^*(\mathcal{M})$. The inner product can also be represented by a second map

$$\mathbf{X} : \mathbf{Y} \rightarrow \mathbb{R}, \quad (0.2.16)$$

$$\mathbf{X} \cdot \mathbf{Y} = X^\mu Y^\nu \mathbf{e}_\mu \cdot \mathbf{e}_\nu = X^\mu Y^\nu g_{\mu\nu}, \quad (0.2.17)$$

which is a new mapping. The inner product also gives the length $|\mathbf{X}|$ or magnitude of any vector $\mathbf{X} \in \mathcal{T}_p(\mathcal{M})$ as,

$$|\mathbf{X}|^2 = \mathbf{X} \cdot \mathbf{X} = g_{\mu\nu} X^\mu X^\nu. \quad (0.2.18)$$

Another way to think of the inner product is that the metric maps a vector \mathbf{X} to an "equivalent" or "dual" co-vector Ξ such that $\mathbf{X} : \Xi = X^\mu \Xi_\mu = g_{\mu\nu} X^\mu X^\nu$. In component form Ξ is

$$\Xi_\mu = g_{\mu\nu} X^\nu; \quad (0.2.19)$$

this use of the metric to map a vector to its corresponding co-vector (and vica versa) is extremely useful. Without loss of information we can write $\Xi_\nu = X_\nu$ to make it obvious that $X_\nu = X^\mu g_{\mu\nu}$ and this convention will be used from now on.

The metric also assigns an inner product and a length measure on the co-tangent plane $\mathcal{T}_p^*(\mathcal{M})$ but instead using the inverse components $g^{\mu\nu} = (g^{-1})_{\mu\nu}$,

$$g^{\mu\nu} = \boldsymbol{\theta}^\mu \cdot \boldsymbol{\theta}^\nu. \quad (0.2.20)$$

Similarly to before the inner product of two co-vectors $\boldsymbol{\omega}$ and $\boldsymbol{\sigma}$ is

$$\boldsymbol{\omega} \cdot \boldsymbol{\sigma} = \omega_\mu \sigma_\nu \boldsymbol{\theta}^\mu \cdot \boldsymbol{\theta}^\nu = \omega_\mu \sigma_\nu g^{\mu\nu} = \omega_\mu \sigma^\mu. \quad (0.2.21)$$

The reason that $g^{\mu\nu}$ must be the inverse matrix to $g_{\mu\nu}$ is as follows. For a vector $x^\mu \mathbf{e}_\mu$ and a co-vector $\omega_\mu \boldsymbol{\theta}^\mu$ we would like,

$$\mathbf{X} : \boldsymbol{\omega} = \mathbf{g}(\mathbf{X}, \star) : \mathbf{g}^{-1}(\boldsymbol{\omega}, \star), \quad (0.2.22)$$

$$X^\mu \omega_\mu = X_\mu \omega^\mu, \quad (0.2.23)$$

$$= X^\rho g_{\rho\mu} g^{\mu\sigma} \omega_\sigma \quad (0.2.24)$$

which is only true if $g_{\rho\mu} g^{\mu\sigma} = \delta_\rho^\sigma$ which is true by definition if $(g^{-1})_{\mu\nu} = g^{\mu\nu}$.

Not only has the metric provided us with an inner product and a length on tangent planes and cotangent planes, but it has also given a mapping between the two. The metric can raise and lower indices on general tensors such as

$$T^{\mu\nu\dots}_{\alpha\beta} = g^{\mu\rho} g_{\beta\sigma} T^\nu_{\rho\alpha}{}^{\sigma\dots}. \quad (0.2.25)$$

0.2.4 Maps Between Manifolds

In section we will be interested in the maps between two manifolds M and N . This has many uses such as pushing and pulling tensors between manifolds, allowing us to calculate a Lie derivative of tensor fields and finding the metric (or any tensor field) on an embedded surface; this very importantly allowed us to perform the 3+1 decomposition [REF] on a spacetime.

Defining a smooth map $\Phi : M \rightarrow N$ between manifolds on some coordinate patch and labelling coordinates $x^\mu \in M$ and $y^\mu \in N$. The map $\Phi : x^\mu \rightarrow y^\mu$ gives $y^\mu = \Phi^\mu(x^\nu)$, or equivalently $y^\mu(x^\nu)$. Scalar functions must also map trivially $f_N(y^\mu(x^\nu)) = f_M(x^\mu)$ where $f_N \in N$ and $f_M \in M$, thus we will no longer identify which manifold a function is on. The map Φ allows us to push the vector $\mathbf{X} \in \mathcal{T}_p(M)$ to $\Phi_* \mathbf{X} \in \mathcal{T}_q(N)$, where $q = \Phi(p)$, in a way such that its action on a function f is the same in either manifold.

$$\mathbf{X}(f) \Big|_p = \Phi_* \mathbf{X}(f) \Big|_q, \quad (0.2.26)$$

$$X^\mu \frac{\partial f}{\partial x^\mu} = (\Phi_* X)^\nu \frac{\partial f}{\partial y^\nu}, \quad (0.2.27)$$

$$\left(X^\mu \frac{\partial y^\nu}{\partial x^\mu} \right) \frac{\partial f}{\partial y^\nu} = (\Phi_* X)^\nu \frac{\partial f}{\partial y^\nu} \quad (0.2.28)$$

and hence the components of the push-forward $\Phi_* \mathbf{X}$ can be read off,

$$(\Phi_* X)^\mu = \frac{\partial y^\mu}{\partial x^\nu} X^\nu. \quad (0.2.29)$$

Given a co-vector field $\omega \in \mathcal{T}_p^*(N)$ we can pull the field back from $\mathcal{T}_p^*(M) \leftarrow \mathcal{T}_q^*(N)$, denoted $\Phi^*\omega$, by demanding that $\Phi^*\omega(\mathbf{X})|_p = \omega(\Phi_*\mathbf{X})|_q$. Evaluating this gives

$$\Phi^*\omega(\mathbf{X})|_p = \omega(\Phi_*\mathbf{X})|_q, \quad (0.2.30)$$

$$(\Phi^*\omega)_\mu X^\mu = \omega_\nu (\Phi_*X)^\nu, \quad (0.2.31)$$

$$(\Phi^*\omega)_\mu X^\mu = \omega_\nu \frac{\partial y^\nu}{\partial x^\mu} X^\mu, \quad (0.2.32)$$

and the components of pull-back $\Phi_*\omega$ can be read off,

$$(\Phi_*\omega)_\mu = \omega_\nu \frac{\partial y^\nu}{\partial x^\mu}. \quad (0.2.33)$$

Considering an $(0, 2)$ tensor $\mathbf{T} \in N$, the pullback $\Phi^*\mathbf{T} \in M$ follows simply from demanding that $\mathbf{T}(\Phi_*\mathbf{X}, \Phi_*\mathbf{Y}) = \Phi^*\mathbf{T}(\mathbf{X}, \mathbf{Y})$ where \mathbf{X} and \mathbf{Y} are vector fields on M . The components of the pull-back of \mathbf{T} are therefore

$$(\Phi^*T)_{\mu\nu} = \frac{\partial y^\rho}{\partial x^\mu} \frac{\partial y^\sigma}{\partial x^\nu} T_{\rho\sigma}, \quad (0.2.34)$$

and the pull-back of a generic $(0, q)$ tensor and the push-forward of a generic $(p, 0)$ tensor can be found similarly.

Diffeomorphisms

So far we have only discussed the mapping $\Phi : M \rightarrow N$ which requires a well behaved $\partial y^\nu / \partial x^\mu$. A diffeomorphism is an isomorphism¹ between smooth manifolds $\Phi : M \rightarrow N$, meaning M and N have the same number of dimensions. Two infinitesimally close points $\{p, p + \delta p\} \in \mathcal{M}$ map to two infinitesimally close points $\{q, q + \delta q\} \in \mathcal{N}$ meaning that open sets are preserved. Given that a diffeomorphism is smooth bijective map then it must be invertible with inverse map $\Phi^{-1} : N \rightarrow M$, and $y^\nu(x^\mu)$ has a smooth inverse $x^\nu(y^\mu)$. When an inverse map Φ^{-1} is defined then the pull-back of $(p, 0)$ tensors from N to M along with the push-forward of $(0, q)$ tensors from M to N is possible. This means it is possible to push or pull generic tensors between M and N in any direction. The tangent spaces associated with $p \in \mathcal{M}$ or $q \in \mathcal{N}$ are therefore also preserved under mapping (aswell as open sets) meaning that local structure on the manifold is preserved under the mapping. Two common examples of diffeomorphisms are coordinate changes and translations.

Projection Mappings

As mentioned, maps between manifolds can be used to determine the metric on an embedded surface. This requires us to consider an m dimensional manifold M with metric $g_{\mu\nu}$ and coordinates x^μ aswell as an embedded n dimensional surface N where $n < m$. We can treat N as a separate n dimensional manifold with metric $h_{\mu\nu}$. As before, we can define a map $\Phi : x^\mu \rightarrow y^\mu$

PROJECTIONS HERE I THINK THEN USE IT IN VOLUMES ON MANIFOLDS TO PROJECT THE METRIC.

0.2.5 Lie Derivatives

We have now discussed the necessary formalism to define the Lie derivative. The Lie derivative at a point is the rate of change of a tensor field with respect to a pull-back from a diffeomorphism Φ mapping infinitesimally close points $p, q \in \mathcal{M}$ like $\Phi : p = x^\mu \rightarrow q = x^\mu + \epsilon \xi^\mu$ for some vector field ξ . Like any

¹An isomorphism is a structure preserving bijective map between sets.

good differential operator, the Lie derivative \mathcal{L}_ξ along ξ (and \mathcal{L}_ζ along ζ) should obey,

$$\mathcal{L}_{a\xi+b\zeta}\mathbf{T} = a\mathcal{L}_\xi\mathbf{T} + b\mathcal{L}_\zeta\mathbf{T}, \quad (0.2.35)$$

$$\mathcal{L}_\xi(a\mathbf{T} + b\mathbf{W}) = a\mathcal{L}_\xi\mathbf{T} + b\mathcal{L}_\xi\mathbf{W}, \quad (0.2.36)$$

$$\mathcal{L}_\xi(f\mathbf{T}) = \mathbf{T}\mathcal{L}_\xi f + f\mathcal{L}_\xi\mathbf{T}, \quad (0.2.37)$$

for constant $\{a, b\}$, function f and generic tensorial objects of same type \mathbf{T} and \mathbf{W} .

The simplest example is the Lie derivative of a scalar field ϕ , denoted $\mathcal{L}_\xi\phi$ with respect to vector field ξ , is

$$\mathcal{L}_\xi\phi = \lim_{\epsilon \rightarrow 0} \left[\frac{\Phi^*\phi|_q - \phi|_p}{\epsilon} \right], \quad (0.2.38)$$

$$= \lim_{\epsilon \rightarrow 0} \left[\frac{\phi(x^\mu + \epsilon\xi^\mu) - \phi(x^\mu)}{\epsilon} \right], \quad (0.2.39)$$

$$= \xi^\mu \partial_\mu \phi, \quad (0.2.40)$$

which reduces to the directional derivative of ϕ with respect to ξ . Next let's calculate the Lie derivative of a vector field \mathbf{X} with respect to vector field ξ . Starting with the same definition as Eq. (0.2.41), and using $y^\mu = x^\mu + \epsilon\xi^\mu$, the Lie derivative of \mathbf{X} is

$$(\mathcal{L}_\xi X)^\mu = \lim_{\epsilon \rightarrow 0} \left[\frac{(\Phi^* X|_q)^\mu - X|_p^\mu}{\epsilon} \right], \quad (0.2.41)$$

$$= \lim_{\epsilon \rightarrow 0} \left[\frac{\frac{\partial x^\mu}{\partial y^\nu} X^\nu(x^\rho + \epsilon\xi^\rho) - X^\mu(x^\rho)}{\epsilon} \right], \quad (0.2.42)$$

$$= \lim_{\epsilon \rightarrow 0} \left[\frac{(\delta_\nu^\mu - \epsilon\partial_\nu\xi^\mu)X^\nu(x^\rho + \epsilon\xi^\rho) - X^\mu(x^\rho)}{\epsilon} \right], \quad (0.2.43)$$

$$= \lim_{\epsilon \rightarrow 0} \left[\frac{-\epsilon\partial_\nu\xi^\mu X^\nu(x^\rho + \epsilon\xi^\rho) + X^\mu(x^\rho + \epsilon\xi^\rho) - X^\mu(x^\rho)}{\epsilon} \right], \quad (0.2.44)$$

$$= \lim_{\epsilon \rightarrow 0} \left[\frac{-\epsilon\partial_\nu\xi^\mu X^\nu(x^\rho) + X^\mu(x^\rho + \epsilon\xi^\rho) - X^\mu(x^\rho) + \mathcal{O}(\epsilon^2)}{\epsilon} \right], \quad (0.2.45)$$

$$= \xi^\nu \partial_\nu X^\mu - X^\nu \partial_\nu \xi^\mu. \quad (0.2.46)$$

The Lie derivative for co-vectors and tensors can be derived in the same way, but can be quickly derived from the Liebnitz rule as follows. Define a scalar field ψ , vector field \mathbf{X} and co-vector field ω , where $\psi = X^\mu \omega_\mu$, then it follows that

$$\mathcal{L}_\xi\psi = \xi^\mu \partial_\mu \psi = X^\nu \xi^\mu \partial_\mu \omega_\nu + \omega_\nu \xi^\mu \partial_\mu X^\nu, \quad (0.2.47)$$

$$= \mathcal{L}_\xi(X^\mu \omega_\mu), \quad (0.2.48)$$

$$= \omega_\mu (\mathcal{L}_\xi X)^\mu + X^\mu (\mathcal{L}_\xi \omega)_\mu, \quad (0.2.49)$$

$$X^\mu (\mathcal{L}_\xi \omega)_\mu = X^\nu \xi^\mu \partial_\mu \omega_\nu + \omega_\nu \xi^\mu \partial_\mu X^\nu - \omega_\mu (\mathcal{L}_\xi X)^\mu, \quad (0.2.50)$$

$$(\mathcal{L}_\xi \omega)_\mu = \xi^\nu \partial_\nu \omega_\mu + \omega_\nu \partial_\mu \xi^\nu. \quad (0.2.51)$$

Derivatives of a generic tensor \mathbf{T} follows simply, for example

$$(\mathcal{L}_\xi T)^{\alpha\beta\dots}_{\mu\nu\dots} = \xi^\sigma \partial_\sigma T^{\alpha\beta\dots}_{\mu\nu\dots} + T^{\alpha\beta\dots}_{\sigma\nu\dots} \partial_\mu \xi^\sigma + T^{\alpha\beta\dots}_{\mu\sigma\dots} \partial_\nu \xi^\sigma + \dots - T^{\sigma\beta\dots}_{\mu\nu\dots} \partial_\sigma \xi^\alpha - T^{\alpha\sigma\dots}_{\mu\nu\dots} \partial_\sigma \xi^\beta - \dots \quad (0.2.52)$$

0.2.6 Lengths on Manifolds

The natural entry point for studying curved geometry is to revisit Pythagoras' theorem. For this we need a manifold \mathcal{M} equipped with a metric g , written as (\mathcal{M}, g) for short. The distance ds between two infinitesimally close points $p \in \mathcal{M}$ and $p + \delta p \in \mathcal{M}$, with coordinates x^μ and $x^\mu + dx^\mu$, is given by

$$ds^2 = \mathbf{g}(\mathbf{dx}, \mathbf{dx}) = g_{\mu\nu} dx^\mu dx^\nu, \quad (0.2.53)$$

where $g_{\mu\nu}$ are the components of the metric tensor. This is the generalisation of Eq. (0.2.2) to curved space; notably the line element can now have varying coefficients from $g_{\mu\nu}$ and cross terms such as $dx^\mu dy^\nu$. The special choice of $g_{\mu\nu} = \delta_{\mu\nu}$ gives us flat space, also called Euclidean space, where $\delta_{\mu\nu} = 1$ if $\mu = \nu$ and vanishes otherwise. With the line element defined, we can immediately apply it to calculating the length of a general curve in curved space. Consider the curve Γ consisting of a set of smoothly connected points $p(\lambda) \in \mathcal{M}$ smoothly parameterised by λ . We can calculate the length Δs of the curve between $\lambda_1 \geq \lambda \geq \lambda_0$ by

$$ds^2 = \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} g_{\mu\nu} d\lambda^2, \quad (0.2.54)$$

$$\Delta s = \int_{\lambda_0}^{\lambda_1} \sqrt{\left(\frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} g_{\mu\nu} \right)} d\lambda. \quad (0.2.55)$$

In the simplified case where λ is one of the coordinates, say ξ , the length Δs becomes,

$$\Delta s = \int_{\xi_0}^{\xi_1} \sqrt{g_{\xi\xi}} d\xi. \quad (0.2.56)$$

0.2.7 Volumes on Manifolds

The natural extension to measuring the length of a curve is the measurement of volumes; of course we still require a metric \mathbf{g} over the manifold \mathcal{M} . Let's say that in a coordinate system x^μ we can define the volume V of some subregion M of \mathcal{M} by integrating some weight function $w(x^\mu)$,

$$V = \int_M w(x^\mu) dx^1 dx^2 \dots dx^n, \quad (0.2.57)$$

over M . To find $w(\mu)$, start by defining an orthogonal coordinate transformation $x^\mu \rightarrow \tilde{x}^\mu$ such that $\tilde{\mathbf{g}}$ is diagonal and $\det(\mathbf{g}) = \det(\tilde{\mathbf{g}})$; this is always possible as \mathbf{g} is real and symmetric. In this coordinate system, the volume δV in an infinitesimal cuboid, with i 'th sidelength $\delta \tilde{x}^i$, is

$$\delta V = \left(\sqrt{\tilde{g}_{11}} \delta \tilde{x}^1 \right) \left(\sqrt{\tilde{g}_{22}} \delta \tilde{x}^2 \right) \dots \left(\sqrt{\tilde{g}_{nn}} \delta \tilde{x}^n \right), \quad (0.2.58)$$

where Eq. (0.2.56) was used to get the length between each \tilde{x}^i and $\tilde{x}^i + \delta \tilde{x}^i$. Given that $\tilde{\mathbf{g}}$ is diagonal we know the i 'th eigenvalue $\tilde{\lambda}_i = \tilde{g}_{ii}$ and therefore $\det(\tilde{\mathbf{g}}) = \prod_i \tilde{g}_{ii}$. Thus the volume δV can be rewritten

$$\delta V = \sqrt{|\det(\tilde{\mathbf{g}})|} \delta \tilde{x}^1 \delta \tilde{x}^2 \dots \delta \tilde{x}^n, \quad (0.2.59)$$

and the formula for the finite volume M is

$$V = \int_M \sqrt{|\det(\tilde{\mathbf{g}})|} d\tilde{x}^1 d\tilde{x}^2 \dots d\tilde{x}^n, \quad (0.2.60)$$

and the form of the weight function in \tilde{x}^μ coordinates is $w(\tilde{x}^\mu) = \sqrt{|\det(\tilde{\mathbf{g}})|}$. We are now free to transform back from $\tilde{x}^\mu \rightarrow x^\mu$, and given that the transformation is orthogonal we know that $\det(\mathbf{g}) = \det(\tilde{\mathbf{g}})$ and the Jacobian matrix \mathbf{J} of the coordinate transformation has $\det(\mathbf{J}) = 1$, therefore

$$V = \int_M \sqrt{|\det(\mathbf{g})|} dx^1 dx^2 \dots dx^n, \quad (0.2.61)$$

which holds for any non-diagonal, real and symmetric metric \mathbf{g} . In general we will now denote the determinant of a metric $\det(\mathbf{g})$ with the lower case letter g . When dealing with a pseudo-Riemannian manifold with a negative determinant, such as spacetime, it is more common to see $\sqrt{-g}$ written rather than $\sqrt{|g|}$ giving

$$V = \int_M \sqrt{-g} dx^1 dx^2 \dots dx^n. \quad (0.2.62)$$

Equation (0.2.60) can also be used to find the volume (or area) of a lower dimensional sub-volume. First cover the new sub-volume A with coordinates y^μ , where $\mu \in \{1, 2, \dots, m\}$ for $m < n$, and then calculate the metric h which can be done using the pullback of g

$$h_{\sigma\rho} = (\Phi^*g)_{\sigma\rho} = \frac{\partial x^\mu}{\partial y^\sigma} \frac{\partial x^\nu}{\partial y^\rho} g_{\mu\nu} \quad (0.2.63)$$

as shown in Section 0.2.4. Using y^μ and h , the volume V_A of A as

$$V_A = \int_A \sqrt{|h|} dy^1 dy^2 \dots dy^m. \quad (0.2.64)$$

NEED TO RE-WRITE THE MAPS SECTION TO HAVE PROJECTIONS AT THE END THEN QUOTE IT FOR THE PROJECTED METRIC HERE INSTEAD OF DERIVING THE PROJECTED METRIC HERE

0.2.8 Geodesics

For a manifold equipped with metric (\mathcal{M}, g) the curve with shortest distance between two points $p, q \in \mathcal{M}$ is called a geodesic. To find the geodesic joining p and q we need to use calculus of variation on the total length Δs from Eq. (0.2.55) of a general curve between two points. Given that the integrand \mathcal{L} of Eq. (0.2.55) is a function like $\mathcal{L}(x^\mu, \dot{x}^\mu)$, where the dot means differentiation by λ , we can use the Euler-Lagrange equation,

$$\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{d}{d\lambda} \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = 0 \quad (0.2.65)$$

to give a differential equation with solution being a geodesic. Applyin the EL equation to the integrand of Eq. (0.2.55) is algebraically messy, it is easier to square the integrand and start from \mathcal{L}^2 giving the same solution if $d\mathcal{L}/d\lambda = 0^2$,

$$\frac{\partial \mathcal{L}^2}{\partial x^\alpha} - \frac{d}{d\lambda} \frac{\partial \mathcal{L}^2}{\partial \dot{x}^\alpha} = 0, \quad (0.2.66)$$

$$\frac{\partial}{\partial x^\alpha} (g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu) - \frac{d}{d\lambda} \frac{\partial}{\partial \dot{x}^\alpha} (g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu) = 0, \quad (0.2.67)$$

$$(\partial_\alpha g_{\mu\nu}) \dot{x}^\mu \dot{x}^\nu - 2 \frac{d}{d\lambda} (g_{\alpha\nu} \dot{x}^\nu) = 0, \quad (0.2.68)$$

$$(\partial_\alpha g_{\mu\nu}) \dot{x}^\mu \dot{x}^\nu - 2 (\dot{x}^\rho \partial_\rho (g_{\alpha\nu}) \dot{x}^\nu) - 2 \ddot{x}^\nu g_{\alpha\nu} = 0. \quad (0.2.69)$$

Rearranging and multiplying by $g^{\alpha\beta}$ gives

$$\ddot{x}^\beta + \frac{1}{2} g^{\alpha\beta} (\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu}) \dot{x}^\mu \dot{x}^\nu = 0, \quad (0.2.70)$$

$$\ddot{x}^\beta + \Gamma_{\mu\nu}^\beta \dot{x}^\mu \dot{x}^\nu = 0, \quad (0.2.71)$$

where $\Gamma_{\mu\nu}^\beta$ is the components of the connection-symbol from Eq. (0.3.77). A trivial solution to Eq. (0.2.71) is in flat space using cartesian coordinates where $\Gamma_{\mu\nu}^\beta = 0$ and therefore $\ddot{x}^\beta = 0$ so \dot{x}^β is a constant; this tells us the shortest distance between two points in flat space is a straight line. In other words, geodesics are straight lines in flat space.

²Given that \mathcal{L} is homogeneneus to degree k , $\dot{x}^i \partial \mathcal{L} / \partial \dot{x}^i = k \mathcal{L}$ for constant k , one can show that $d\mathcal{L}/d\lambda = 0$ if the Euler-Lagrange equation is assumed.

Non-affine Geodesics

The equation of a geodesic given above is true for an affinely parameterised curve. An affine parameter λ is defined so that the length of a curve Δs between two parameter values λ_0 and λ_1 is given by $\Delta s = k(\lambda_1 - \lambda_0)$ for constant k ; the arclength along a curve is linearly proportional to the value of the λ . The reason this happened with the calculation above is that $d\mathcal{L}/d\lambda = 0$ was assumed.

A non-affine parameter μ could equally be used to describe the curve. Writing $\mu(\lambda)$ the geodesic equation is transformed as shown,

$$\frac{d^2 x^\beta}{d\lambda^2} + \Gamma_{\mu\nu}^\beta \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0, \quad (0.2.72)$$

$$\left(\frac{d^2 \mu}{d\lambda^2} \frac{d}{d\mu} + \left(\frac{d\mu}{d\lambda} \right)^2 \frac{d^2}{d\mu^2} \right) x^\beta + \Gamma_{\mu\nu}^\beta \frac{dx^\mu}{d\mu} \frac{dx^\nu}{d\mu} \left(\frac{d\mu}{d\lambda} \right)^2 = 0, \quad (0.2.73)$$

$$\frac{d^2 x^\beta}{d\mu^2} + \Gamma_{\mu\nu}^\beta \frac{dx^\mu}{d\mu} \frac{dx^\nu}{d\mu} = - \left(\frac{d\mu}{d\lambda} \right)^{-2} \frac{d^2 \mu}{d\lambda^2} \frac{dx^\beta}{d\mu}, \quad (0.2.74)$$

$$\frac{d^2 x^\beta}{d\mu^2} + \Gamma_{\mu\nu}^\beta \frac{dx^\mu}{d\mu} \frac{dx^\nu}{d\mu} = -f(\mu) \frac{dx^\beta}{d\mu}, \quad (0.2.75)$$

$$(0.2.76)$$

which is the same as the affine geodesic equation except with an extra non-zero right hand side proportional to $dx^\beta/d\mu$ and some function $f(\mu)$.

IS THIS WELL ENOUGH EXPLAINED, ESPECIALLY THE DLDLAMBDA=0 Bit?

0.3 Tensor Calculus and Curvature

0.3.1 General Covariance and Coordinate transformations

Many laws of physics can be expressed as tensor field equations where a tensor field is the assignment of a tensor to each point in space. This assignment must be smooth as it is to describe physical quantities. The power of tensor algebra and tensor calculus is that if a tensor field equation can be written in one coordinate system then must hold (in index form) in all sensible coordinate system. This is a consequence of the tensor transformation law. Looking back, we can write a generic vector \mathbf{X} as $X^\mu e_\mu = X^\mu \partial_\mu$ and if we choose a coordinate transformation $x^\mu \rightarrow \tilde{x}^\mu$ then we see that in the transformed coordinate system the vector field \mathbf{X} , written $\tilde{\mathbf{X}}$, becomes

$$\tilde{\mathbf{X}} = \tilde{X}^\mu \frac{\partial}{\partial \tilde{x}^\mu}, \quad (0.3.1)$$

$$= \tilde{X}^\mu \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\nu}, \quad (0.3.2)$$

$$= X^\nu \frac{\partial}{\partial x^\nu}, \quad (0.3.3)$$

$$= \mathbf{X}, \quad (0.3.4)$$

where the components $X^\nu = \tilde{X}^\mu \frac{\partial x^\nu}{\partial \tilde{x}^\mu}$ are required to transform in order to ensure $\mathbf{X} = \tilde{\mathbf{X}}$. This says that the underlying geometric object (a vector in this case) is independent of the coordinates used to describe them; the tradeoff for this useful property is that the vectors components X^μ have to transform under the tensor transformation law, effectively opposing the transformation of the basis vectors. Working from a co-vector ω we can write it as $\omega_\mu \theta^\mu = \omega_\mu dx^\mu$ in component-basis form [REF THIS?] and the

same coordinate transform gives

$$\tilde{\omega} = \tilde{\omega}_\mu d\tilde{x}^\mu, \quad (0.3.5)$$

$$= \tilde{\omega}_\mu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} dx^\nu, \quad (0.3.6)$$

$$= \omega_\nu dx^\nu, \quad (0.3.7)$$

where the co-vector components transform like $\omega_\nu = \tilde{\omega}_\mu \frac{\partial \tilde{x}^\mu}{\partial x^\nu}$, the opposite way to the vector components. These transformation laws ensure that a scalar field created from the product of a vector field and a co-vector field, like $\omega : \mathbf{X}$, is a Lorentz scalar not transforming under coordinate transformations. This can be seen from

$$\tilde{\omega} : \tilde{\mathbf{X}} = \tilde{X}^\mu \tilde{\omega}_\mu, \quad (0.3.8)$$

$$= X^\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \omega_\rho, \quad (0.3.9)$$

$$= X^\nu \frac{\partial x^\rho}{\partial x^\nu} \omega_\rho, \quad (0.3.10)$$

$$= X^\nu \delta_\nu^\rho \omega_\rho, \quad (0.3.11)$$

$$= X^\nu \omega_\nu, \quad (0.3.12)$$

$$= \omega : \mathbf{X}. \quad (0.3.13)$$

The general tensor transformation law can be derived easily from chaining multiple of the previous examples together, for example

$$\tilde{T}^{\mu\nu\dots}_{\rho\sigma\dots} = T^{\alpha\beta\dots}_{\gamma\delta\dots} \left(\frac{\partial \tilde{x}^\mu}{\partial x^\alpha} \frac{\partial \tilde{x}^\nu}{\partial x^\beta}, \dots \times \frac{\partial x^\gamma}{\partial \tilde{x}^\rho} \frac{\partial x^\delta}{\partial \tilde{x}^\sigma}, \dots \right). \quad (0.3.14)$$

Tensor Densities

A tensor density is the generalisation of a tensor field obeying the tensor transformation law in Eq. (0.3.14) to a tensor field multiplied by a power of the determinant of the Jacobean matrix of a coordinate transformation. One important example of a tensor density is the volume element $\sqrt{-g}$, this is a scalar density. This object does not have any indices so at first glance may pass for a true scalar field. However, when a coordinate transformation $x^\mu \rightarrow \tilde{x}^\mu$ is applied we find that $\sqrt{-g} \rightarrow \sqrt{-\tilde{g}} \neq \sqrt{-g}$ but for a general scalar field ϕ we find $\phi \rightarrow \tilde{\phi} = \phi$; therefore $\sqrt{-g}$ cannot be a scalar field. This can be shown explicitly by look at the the determinant of the metric,

$$\sqrt{-\tilde{g}} = \sqrt{\det(-\tilde{g}_{\mu\nu})}, \quad (0.3.15)$$

$$= \sqrt{\det \left(-g_{\alpha\beta} \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\beta}{\partial \tilde{x}^\nu} \right)}, \quad (0.3.16)$$

$$= \sqrt{-g} \det \left(\frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \right), \quad (0.3.17)$$

and as can be seen, the volume element picks up a factor of the determinant of the Jacobean. This property shows up in multidimensional intergrals, for instance when taking a three-dimensional integral over cartesian coordinates we must replace the $dx dy dz \rightarrow dr d\theta d\phi r^2 \sin(\theta)$ [MAYBE DO THIS LATER AFTER INTEGRATION ON MANIFODS?]

A tensor density \mathcal{T} of weight w can be written in the form,

$$\mathcal{T} = \sqrt{-g}^w \mathbf{T}, \quad (0.3.18)$$

where \mathbf{T} is a tensor obeying the tensor transformation law. It should be noted that a tensor density of weight zero is a regular tensors and the weight of a tensor has nothing to do with the rank of the tensor.

Lie Derivatives of Tensor Densities

To calculate the Lie derivative of a tensor density, first the Lie derivative of $\sqrt{-g}$ should be calculated, and in order to calculate the Lie derivative of the volume element a preliminary result is needed. Following the definition of a Lie derivative in section 0.2.5 and setting $y^\mu = x^\mu + \epsilon \xi^\mu$, the determinant of the Jacobean matrix is,

$$\det \left(\frac{\partial x^\mu}{\partial y^\rho} \right) = \det \left(\frac{\partial y^\mu}{\partial x^\rho} \right)^{-1}, \quad (0.3.19)$$

$$= \det \left(\delta_\rho^\mu + \epsilon \frac{\partial \xi^\mu}{\partial x^\rho} \right)^{-1}, \quad (0.3.20)$$

$$= \det \begin{pmatrix} 1 + \epsilon \frac{\partial \xi^1}{\partial x^1} & \epsilon \frac{\partial \xi^1}{\partial x^2} & \epsilon \frac{\partial \xi^1}{\partial x^3} & \epsilon \frac{\partial \xi^1}{\partial x^4} \\ \epsilon \frac{\partial \xi^2}{\partial x^1} & 1 + \epsilon \frac{\partial \xi^2}{\partial x^2} & \epsilon \frac{\partial \xi^2}{\partial x^3} & \epsilon \frac{\partial \xi^2}{\partial x^4} \\ \epsilon \frac{\partial \xi^3}{\partial x^1} & \epsilon \frac{\partial \xi^3}{\partial x^2} & 1 + \epsilon \frac{\partial \xi^3}{\partial x^3} & \epsilon \frac{\partial \xi^3}{\partial x^4} \\ \epsilon \frac{\partial \xi^4}{\partial x^1} & \epsilon \frac{\partial \xi^4}{\partial x^2} & \epsilon \frac{\partial \xi^4}{\partial x^3} & 1 + \epsilon \frac{\partial \xi^4}{\partial x^4} \end{pmatrix}^{-1}, \quad (0.3.21)$$

$$= \left(1 + \epsilon \sum_i \frac{\partial \xi^i}{\partial x^i} + \mathcal{O}(\epsilon^2) \right)^{-1}, \quad (0.3.22)$$

$$= 1 - \epsilon \partial_\mu \xi^\mu + \mathcal{O}(\epsilon^2), \quad (0.3.23)$$

where four dimensions was used for clarity, but the calculation works exactly the same in any number of dimensions. Using this result and the definition of a Lie derivative, $\mathcal{L}_\xi \sqrt{-g}$ evaluates to

$$\mathcal{L}_\xi \sqrt{-g} = \lim_{\epsilon \rightarrow 0} \left(\frac{\Phi^* \sqrt{-g}|_q - \sqrt{-g}|_p}{\epsilon} \right), \quad (0.3.24)$$

$$= \lim_{\epsilon \rightarrow 0} \left(\frac{\sqrt{-\det \left(g_{\mu\nu}(y^\alpha) \frac{\partial x^\mu}{\partial y^\rho} \frac{\partial x^\nu}{\partial y^\sigma} \right)} - \sqrt{-g}(x^\alpha)}{\epsilon} \right), \quad (0.3.25)$$

$$= \lim_{\epsilon \rightarrow 0} \left(\frac{\sqrt{-g}(y^\alpha) \det \left(\frac{\partial x^\mu}{\partial y^\rho} \right) - \sqrt{-g}(x^\alpha)}{\epsilon} \right), \quad (0.3.26)$$

$$= \lim_{\epsilon \rightarrow 0} \left(\frac{\sqrt{-g}(y^\alpha) (1 - \epsilon \partial_\mu \xi^\mu + \mathcal{O}(\epsilon^2)) - \sqrt{-g}(x^\alpha)}{\epsilon} \right), \quad (0.3.27)$$

$$= \lim_{\epsilon \rightarrow 0} \left(\frac{[\sqrt{-g}(x^\alpha) + \epsilon \xi^\mu \partial_\mu \sqrt{-g}(x^\alpha)] (1 - \epsilon \partial_\mu \xi^\mu + \mathcal{O}(\epsilon^2)) - \sqrt{-g}(x^\alpha)}{\epsilon} \right), \quad (0.3.28)$$

$$= \lim_{\epsilon \rightarrow 0} \left(\frac{\epsilon \xi^\mu \partial_\mu \sqrt{-g}(x^\alpha) - \epsilon \sqrt{-g}(x^\alpha) \partial_\mu \xi^\mu + \mathcal{O}(\epsilon^2)}{\epsilon} \right), \quad (0.3.29)$$

$$\mathcal{L}_\xi \sqrt{-g} = \xi^\mu \partial_\mu \sqrt{-g} - \sqrt{-g} \partial_\mu \xi^\mu. \quad (0.3.30)$$

Given that $\mathcal{L}_\xi \sqrt{-g}$ has been calculated, it is straightforward to calculate the Lie derivative of a tensor density $\mathcal{T} = \sqrt{-g}^w \mathbf{T}$ of weight w , where \mathbf{T} is a regular tensor, using the Leibnitz rule in Eq. (0.2.37). The Lie derivative is,

$$\mathcal{L}_\xi \mathcal{T} = \sqrt{-g}^w \left(\mathcal{L}_\xi \mathbf{T} - w \mathbf{T} \partial_\mu \xi^\mu + w \frac{1}{\sqrt{-g}} \xi^\mu \partial_\mu \sqrt{-g} \right), \quad (0.3.31)$$

where the Leibnitz rule of differentiation has been used. As would be expected, setting $w = 0$ returns the regular Lie derivative of a tensor.

CHECK THIS IS CORRECT AND USED CORRECTLY LATER IN THE CCZ4 EQUATIONS SECTION

0.3.2 The Covariant Derivative

There are many types of derivative on a manifold, all related to each other, that we care about in the context of General Relativity. The simplest is the partial derivative, denoted

$$\partial_\mu = \frac{\partial}{\partial x^\mu}, \quad (0.3.32)$$

which works much the same as always. Two other derivatives are the exterior derivative [MAYBE REF THIS] and the Lie derivative which are discussed in [REF] and section 0.2.5 respectively.

The purpose of the covariant derivative, denoted ∇_μ , is to generalise the partial derivative to curved spaces (or curvilinear coordinates). The covariant derivative exactly reduces to the partial derivative (∂_μ) in flat space with cartesian coordinates. As suggested by the name, the covariant derivative of an object is covariant; it obeys the tensor transformation law in section 0.3.1. The covariant derivative uses a vector field \mathbf{X} to map a (p, q) tensor field \mathbf{T} to a (p, q) tensor field $\nabla_{\mathbf{X}}\mathbf{T}$. Requiring the covariant derivative of a tensor to return another tensor may sound pedantic but it allows the writing of physical differential equations that are covariant, i.e. that hold in all coordiante systems. Encoding the laws of physics with tensor differential equations is explored in more detail in Section [REF!].

To be the analogue of the partial derivative, three properties of the covariant derivative are required,

$$\nabla_{f\mathbf{X}+g\mathbf{Y}}\mathbf{T} = f\nabla_{\mathbf{X}}\mathbf{T} + g\nabla_{\mathbf{Y}}\mathbf{T}, \quad (0.3.33)$$

$$\nabla_{\mathbf{X}}(a\mathbf{T} + b\mathbf{W}) = a\nabla_{\mathbf{X}}\mathbf{T} + b\nabla_{\mathbf{X}}\mathbf{W}, \quad (0.3.34)$$

$$\nabla_{\mathbf{X}}(f\mathbf{T}) = f\nabla_{\mathbf{X}}\mathbf{T} + \mathbf{T}\nabla_{\mathbf{X}}f, \quad (0.3.35)$$

where a and b are constants, f and g are functions and \mathbf{T} and \mathbf{W} are tensors of equal type. These are exactly the same as the conditions imposed on Lie derivatives in section 0.2.5. From now on the coraviant derivative $\nabla_{\mathbf{e}_\mu}$ with respect to a basis vector \mathbf{e}_μ will be written as ∇_μ .

Lets start by finding the covariant derivative of a of a scalar field φ . The partial derivative $\partial_\mu\varphi$ obeys the tensor transformation law for a co-vector,

$$\frac{\partial}{\partial \tilde{x}^\mu} \tilde{\varphi} = \frac{\partial}{\partial \tilde{x}^\mu} \varphi, \quad (0.3.36)$$

$$= \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\nu} \varphi, \quad (0.3.37)$$

and therefore the $\nabla_\mu\varphi = \partial_\mu\varphi$. Note that $\varphi = \tilde{\varphi}$ for any point p as a scalar remains unchanged in a coordinate transformation. Complications arise when taking the partial derivative of any other higher rank tensor; let's demonstrate this with a vector \mathbf{X} .

$$\frac{\partial}{\partial \tilde{x}^\mu} \tilde{X}^\alpha = \frac{\partial}{\partial \tilde{x}^\mu} \left(\frac{\partial \tilde{x}^\alpha}{\partial x^\beta} X^\beta \right), \quad (0.3.38)$$

$$= \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\nu} \left(\frac{\partial \tilde{x}^\alpha}{\partial x^\beta} X^\beta \right), \quad (0.3.39)$$

$$= \underbrace{\frac{\partial \tilde{x}^\alpha}{\partial x^\beta} \frac{\partial x^\nu}{\partial \tilde{x}^\mu}}_{\text{Tensor transformation law}} \frac{\partial}{\partial x^\nu} X^\beta + X^\beta \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\nu} \left(\frac{\partial \tilde{x}^\alpha}{\partial x^\beta} \right), \quad (0.3.40)$$

and as can be seen, only the first term on the right hand side should exist if the components $\partial_\mu X^\alpha$ were to obey the tensor transformation law.

The problem with performing differentiation on tensors is that it requires the comparison of tensors at two different (infinitessimally close) tangent spaces. Lie derivatives circumvented this problem by comparing two two tangent planes with a pullback defined by a diffeomorphism. Another way of overcoming

this problem is to consider how the coordinate basis vectors change over the manifold, not just the components. Defining the covariant derivative of the basis vector as

$$\nabla_{\mathbf{e}_\rho} \mathbf{e}_\nu = \nabla_\rho \mathbf{e}_\nu := \Gamma^\mu_{\nu\rho} \mathbf{e}_\mu \quad (0.3.41)$$

where $\Gamma^\mu_{\nu\rho}$ is called the connection due to it defining a connection between neighbouring tangent space. The connection can be used to get the covariant derivative of the vector field $\mathbf{X} = X^\rho \mathbf{e}_\rho$,

$$\nabla_\rho (X^\nu \mathbf{e}_\nu) = (\partial_\rho X^\nu) \mathbf{e}_\nu + X^\nu (\nabla_\rho \mathbf{e}_\nu), \quad (0.3.42)$$

$$= (\partial_\rho X^\nu) \mathbf{e}_\nu + X^\nu \Gamma^\mu_{\nu\rho} \mathbf{e}_\mu, \quad (0.3.43)$$

$$= (\partial_\rho X^\mu + \Gamma^\mu_{\nu\rho} X^\nu) \mathbf{e}_\mu. \quad (0.3.44)$$

Note that on the first line above we used $\nabla_\rho X^\nu = \partial_\rho X^\nu$ as the X^ν are being treated as a set of scalar function coefficients multiplying the basis vectors \mathbf{e}_μ . Strictly we should write the covariant derivative of \mathbf{X} as

$$\nabla \mathbf{X} = (\nabla X)_\sigma^\mu \mathbf{e}_\mu \otimes \boldsymbol{\theta}^\sigma \quad (0.3.45)$$

but for convenience the coefficients $(\nabla X)_\sigma^\mu$ are usually denoted as,

$$\nabla_\sigma X^\mu = \partial_\sigma X^\mu + \Gamma^\mu_{\nu\sigma} X^\nu. \quad (0.3.46)$$

This is a slight abuse of notation as $\nabla_\sigma X^\mu$ might be understood as the covariant derivative of the components X^μ , but really it denotes the component $\boldsymbol{\theta}^\sigma \cdot (\nabla \mathbf{X}) \cdot \mathbf{e}_\mu$ where $\nabla \mathbf{X}$ is given in Eq. (0.3.45).

Given that the covariant derivative of a scalar reduces to the partial derivative we can see that

$$\nabla_\rho (\mathbf{e}_\mu : \boldsymbol{\theta}^\nu) = 0, \quad (0.3.47)$$

and using the Liebnitz rule we see that

$$\nabla_\rho (\mathbf{e}_\nu : \boldsymbol{\theta}^\mu) = (\nabla_\rho \mathbf{e}_\nu) : \boldsymbol{\theta}^\mu + \mathbf{e}_\nu : (\nabla_\rho \boldsymbol{\theta}^\mu), \quad (0.3.48)$$

$$= (\Gamma^\sigma_{\nu\rho} \mathbf{e}_\sigma) : \boldsymbol{\theta}^\mu + \mathbf{e}_\nu : (\nabla_\rho \boldsymbol{\theta}^\mu), \quad (0.3.49)$$

$$= \Gamma^\mu_{\nu\rho} + \mathbf{e}_\nu : (\nabla_\rho \boldsymbol{\theta}^\mu), \quad (0.3.50)$$

$$\mathbf{e}_\nu : (\nabla_\rho \boldsymbol{\theta}^\mu) = -\Gamma^\mu_{\nu\rho} \quad (0.3.51)$$

therefore we must have

$$\nabla_\rho \boldsymbol{\theta}^\mu = -\Gamma^\mu_{\nu\rho} \boldsymbol{\theta}^\nu. \quad (0.3.52)$$

In an identical way to before, we might ask what is the covariant derivative of a co-vector $\boldsymbol{\omega} = \omega_\alpha \boldsymbol{\theta}^\alpha$. The covariant derivative $\nabla \boldsymbol{\omega}$ can be found like

$$\nabla_\sigma \boldsymbol{\omega} = \nabla_\sigma (\omega_\alpha \boldsymbol{\theta}^\alpha), \quad (0.3.53)$$

$$= \partial_\sigma (\omega_\alpha) \boldsymbol{\theta}^\alpha + \omega_\alpha \nabla_\sigma \boldsymbol{\theta}^\alpha, \quad (0.3.54)$$

$$= \partial_\sigma (\omega_\alpha) \boldsymbol{\theta}^\alpha - \omega_\alpha \Gamma^\alpha_{\nu\sigma} \boldsymbol{\theta}^\nu, \quad (0.3.55)$$

$$= (\partial_\sigma \omega_\alpha - \omega_\nu \Gamma^\nu_{\alpha\sigma}) \boldsymbol{\theta}^\alpha. \quad (0.3.56)$$

Again, we used $\nabla_\sigma \omega_\alpha = \partial_\sigma \omega_\alpha$ as the components ω_α are scalar coefficients of the basis co-vectors $\boldsymbol{\theta}^\alpha$. Similarly to earlier, from now on the components $(\nabla \omega)_{\sigma\alpha}$ are written as $\nabla_\sigma \omega_\alpha$ even though this is a mild abuse of notation.

The covariant derivative of a general tensor can be found by following the simple rule of adding a connection symbol term for each index, for example

$$\nabla_\mu T^{\alpha\beta\dots}_{\lambda\nu\dots} = \partial_\mu T^{\alpha\beta\dots}_{\lambda\nu\dots} + \Gamma^\alpha_{\sigma\mu} T^{\sigma\beta\dots}_{\lambda\nu\dots} + \Gamma^\beta_{\sigma\mu} T^{\alpha\sigma\dots}_{\lambda\nu\dots} + \dots - \Gamma^\sigma_{\lambda\mu} T^{\alpha\beta\dots}_{\sigma\nu\dots} - \Gamma^\sigma_{\nu\mu} T^{\alpha\beta\dots}_{\lambda\sigma\dots} - \dots \quad (0.3.57)$$

COV DERIV OF TENSOR DENSITIES - MAYBE TENSOR DENSITIES ARE THEIR OWN SECTION. COULD BE QUICK, I ASSUME $\text{NABLA}_M \text{UROOTG} = 0$

0.3.3 The Connection

In flat space we are used to the idea that the partial derivative commutes, i.e. $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$, and this is trivially true in curved space too. However, the covariant derivative does not generally commute, $\nabla_\mu \nabla_\nu \neq \nabla_\nu \nabla_\mu$. Applying $\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu$ to a scalar field φ gives,

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu)\varphi = \nabla_\mu \nabla_\nu \varphi - \nabla_\nu \nabla_\mu \varphi, \quad (0.3.58)$$

$$= \nabla_\mu \partial_\nu \varphi - \nabla_\nu \partial_\mu \varphi, \quad (0.3.59)$$

$$= (\partial_\mu \partial_\nu - \partial_\nu \partial_\mu)\varphi - \Gamma_{\nu\mu}^\sigma \partial_\sigma \varphi + \Gamma_{\mu\nu}^\sigma \partial_\sigma \varphi, \quad (0.3.60)$$

$$= (\Gamma_{\mu\nu}^\sigma - \Gamma_{\nu\mu}^\sigma) \partial_\sigma \varphi, \quad (0.3.61)$$

where we used the fact that $\nabla_\mu \varphi = \partial_\mu \varphi$ for a scalar field. This non-commutativity of derivatives on a scalar field is known as torsion.

Torsion

If a connection is torsion free then $(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu)\varphi = 0$ which implies $\Gamma_{\nu\mu}^\sigma = \Gamma_{\mu\nu}^\sigma$. This leads to two important tensor identities; first the antisymmetric derivative of a co-vector,

$$\nabla_\mu A_\nu - \nabla_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu - \underbrace{(\Gamma_{\nu\mu}^\sigma - \Gamma_{\mu\nu}^\sigma)}_{=0} A_\sigma, \quad (0.3.62)$$

$$\nabla_\mu A_\nu - \nabla_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (0.3.63)$$

and the second identity,

$$\nabla_X Y - \nabla_Y X = (X^\mu \nabla_\mu Y^\nu - Y^\mu \nabla_\mu X^\nu) \mathbf{e}_\nu, \quad (0.3.64)$$

$$= (X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu + \Gamma_{\sigma\mu}^\nu X^\mu Y^\sigma - \Gamma_{\sigma\mu}^\nu Y^\mu X^\sigma) \mathbf{e}_\nu, \quad (0.3.65)$$

$$= (X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu + (\Gamma_{\sigma\mu}^\nu - \Gamma_{\mu\sigma}^\nu) X^\mu Y^\sigma) \mathbf{e}_\nu, \quad (0.3.66)$$

$$= (X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu) \mathbf{e}_\nu, \quad (0.3.67)$$

$$\nabla_X Y - \nabla_Y X = [X, Y]. \quad (0.3.68)$$

The commutator bracket $[X, Y]$ of two vectors X and Y is defined by

$$[X^\mu \partial_\mu, Y^\nu \partial_\nu] = X^\mu \partial_\mu (Y^\nu \partial_\nu) - Y^\nu \partial_\nu (X^\mu \partial_\mu), \quad (0.3.69)$$

$$= X^\mu \partial_\mu Y^\nu - Y^\nu \partial_\nu (X^\mu) \partial_\mu + X^\mu Y^\nu \partial_\mu \partial_\nu - Y^\nu X^\mu \partial_\nu \partial_\mu, \quad (0.3.70)$$

$$= (X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu) \partial_\nu, \quad (0.3.71)$$

where the basis vector \mathbf{e}_μ has been written as ∂_μ with the intention of acting on a function f over the manifold \mathcal{M} .

Metric Compatibility

Another useful property that can be imposed on the connection is metric compatibility; this is $\nabla_\mu g_{\rho\sigma} = 0$, where g is the metric tensor, which immediately tells us $\nabla_\mu g^{\alpha\beta} = 0$ as

$$\nabla_\mu \delta_\rho^\alpha = 0 \quad (0.3.72)$$

$$= \nabla_\mu (g^{\alpha\nu} g_{\nu\rho}), \quad (0.3.73)$$

$$= g_{\nu\rho} \nabla_\mu g^{\alpha\nu} + g^{\alpha\nu} \underbrace{\nabla_\mu g_{\nu\rho}}_{=0}, \quad (0.3.74)$$

which implies that $\nabla_\mu g^{\alpha\nu} = 0$. Demanding metric compatibility may seem a little arbitrary but turns out to have many nice algebraic properties such as the raising and lowering of indices with the metric commuting with the covariant derivatives,

$$\nabla_\mu T^{\alpha\beta\dots} = \nabla_\mu g^{\alpha\rho} T_\rho^{\beta\dots} = g^{\alpha\rho} \nabla_\mu T_\rho^{\beta\dots} \quad (0.3.75)$$

and the derivative of a vector's length,

$$\nabla_\alpha(X^\mu X_\mu) = \nabla_\alpha(g_{\mu\nu}X^\mu X^\nu) = 2X^\mu \nabla_\alpha X_\mu = 2X_\mu \nabla_\alpha X^\mu. \quad (0.3.76)$$

The Levi Civita-Connection

A connection that obeys Eqs. (0.3.33, 0.3.34 & 0.3.35), is both torsion-free and metric-compatible, can be demanded; this connection is called the Levi-Civita connection. The Levi-Civita connection will always be assumed from now and leads to a unique choice of connection coefficients,

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}g^{\rho\sigma}(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\nu\mu}), \quad (0.3.77)$$

which are also called Christoffel symbols of the second kind. It is also common to see the connection coefficients with the lowered index

$$\Gamma_{\sigma\mu\nu} = \frac{1}{2}(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\nu\mu}), \quad (0.3.78)$$

this is also called a Christoffel symbol of the first kind. It is very important to note that even though the connection symbols may look like a tensor they are not a tensor. This can easily be seen from applying the tensor transformation law to the Christoffel symbol of the first kind,

$$2\tilde{\Gamma}_{\sigma\mu\nu} = \tilde{\partial}_\mu \tilde{g}_{\sigma\nu} + \tilde{\partial}_\nu \tilde{g}_{\mu\sigma} - \tilde{\partial}_\sigma \tilde{g}_{\nu\mu}, \quad (0.3.79)$$

$$= \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \partial_\alpha \left(\frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} g_{\beta\gamma} \right) + \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} \partial_\gamma \left(\frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} g_{\alpha\beta} \right) - \frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \partial_\beta \left(\frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} g_{\gamma\alpha} \right), \quad (0.3.80)$$

$$= \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} (\partial_\alpha g_{\beta\gamma} + \partial_\gamma g_{\alpha\beta} - \partial_\beta g_{\gamma\alpha}) \\ + \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} g_{\beta\gamma} \partial_\alpha \left(\frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} \right) + \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} g_{\alpha\beta} \partial_\gamma \left(\frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \right) - \frac{\partial x^\beta}{\partial \tilde{x}^\sigma} g_{\gamma\alpha} \partial_\beta \left(\frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} \right), \quad (0.3.81)$$

$$= 2 \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} \Gamma_{\beta\alpha\gamma} \\ + \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} g_{\beta\gamma} \partial_\alpha \left(\frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} \right) + \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} g_{\alpha\beta} \partial_\gamma \left(\frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \right) - \frac{\partial x^\beta}{\partial \tilde{x}^\sigma} g_{\gamma\alpha} \partial_\beta \left(\frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} \right), \quad (0.3.82)$$

$$= 2 \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\beta}{\partial \tilde{x}^\sigma} \frac{\partial x^\gamma}{\partial \tilde{x}^\nu} \Gamma_{\beta\alpha\gamma} + \Xi_{\sigma\mu\nu}. \quad (0.3.83)$$

$$(0.3.84)$$

As can be seen above, if $\Xi_{\sigma\mu\nu} = 0$ then the $\Gamma_{\sigma\mu\nu}$ would transform as an $(0,3)$ tensor, however this is not the case and the existence of nonzero $\Xi_{\sigma\mu\nu}$ means the Christoffel symbol of first kind is not a tensor but instead a symbol. The non-tensor nature of the Christoffel symbol of first kind is sufficient to prove that the Christoffel symbol of second kind is also not a tensor.

Lie Derivatives in the Levi Civita Connection

The covariant derivative ∇ , described in section 0.3.2 can replace all partial derivatives in a Lie derivative due to the connection symbols $(\Gamma^\mu_{\rho\sigma})$ cancelling out.

A special example of a Lie derivative is of the metric tensor, g , giving

$$(\mathcal{L}_\xi g)_{\mu\nu} = \xi^\rho \partial_\rho g_{\mu\nu} + g_{\rho\nu} \partial_\mu \xi^\rho + g_{\mu\rho} \partial_\nu \xi^\rho, \quad (0.3.85)$$

$$= \xi^\rho \nabla_\rho g_{\mu\nu} + g_{\rho\nu} \nabla_\mu \xi^\rho + g_{\mu\rho} \nabla_\nu \xi^\rho, \quad (0.3.86)$$

$$(0.3.87)$$

where $\nabla_\rho g_{\mu\nu} = 0$ is assumed from metric compatibility, described in section 0.3.3. In the case the Lie derivative vanishes we get Killing's equation

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0 \quad (0.3.88)$$

and a vector field ξ satisfying Killing's equation is called a Killing vector.

MAYBE DERIVE THE FORM OF BIG GAMMA?

Normal Coordinates

Consider the set of all affinely parameterised geodesics, parameterised with λ , passing through the point p on a manifold \mathcal{M} . At p , each geodesic has a tangent vector $X^\mu|_p$. For all geodesics, define p to be the origin with $\lambda = 0$ and $x^\mu = 0$. Following a geodesic associated with $X^\mu|_p$ to a parameter value of λ will map to a new point q close to p for small enough λ . Normal coordinates $x^\mu(\lambda)$ at $q(\lambda)$ are defined such that $x^\mu(\lambda) = \lambda X^\mu|_p$. Given that $X^\mu|_p$ is constant, the geodesic equation, from Eq. (0.2.71), becomes

$$\frac{d^2 x^\mu(\lambda)}{d\lambda^2} + \Gamma^\alpha_{\mu\nu} \frac{dx^\mu(\lambda)}{d\lambda} \frac{dx^\nu(\lambda)}{d\lambda} = \Gamma^\alpha_{\mu\nu} X^\mu|_p X^\nu|_p = 0. \quad (0.3.89)$$

Using the Levi-civita connection, the connection symbol must be symmetric in lower two indices and this implies $\Gamma^\alpha_{\mu\nu} = 0$ as the geodesic equation must hold for generic $X^\mu|_p$. From the definition of the covariant derivative,

$$\partial_\alpha g_{\mu\nu} = \nabla_\alpha g_{\mu\nu} + \Gamma^\beta_{\alpha\nu} g_{\mu\beta} + \Gamma^\beta_{\alpha\mu} g_{\beta\nu}, \quad (0.3.90)$$

which must vanish in normal coordinates as $\Gamma^\alpha_{\mu\nu} = 0$ and the Levi-Civita connection demands $\nabla_\alpha g_{\mu\nu} = 0$. Therefore, it is possible to construct a coordinate system that at one point p both $\partial_\alpha g_{\mu\nu} = 0$ and $\Gamma^\alpha_{\mu\nu} = 0$; importantly $\partial_\alpha \partial_\beta g_{\mu\nu} \neq 0$. As well as the metric's first derivative vanishing at p , it is also possible to demand that $g_{\mu\nu} p = \eta_{\mu\nu}$, the Minkowski metric. This can be done with a set of ${}^{(j)}X^\mu|_p$, one for each normal coordinate x^j , that satisfy

$$g({}^{(j)}X|_p, {}^{(k)}X|_p) = \eta_{jk}. \quad (0.3.91)$$

0.3.4 Curvature Tensors

We have already seen that with the Levi-Civita connection, the commuted derivative of a scalar field vanishes. But taking the commuted derivative of a vector field X gives,

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) X^\sigma = (\partial_\mu \nabla_\nu - \partial_\nu \nabla_\mu) X^\sigma + (\Gamma^\sigma_{\mu\rho} \nabla_\nu - \Gamma^\sigma_{\nu\rho} \nabla_\mu) X^\rho - \underbrace{(\Gamma^\rho_{\nu\mu} - \Gamma^\rho_{\mu\nu})}_{=0} \nabla_\rho X^\sigma, \quad (0.3.92)$$

$$= \underbrace{(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu)}_{=0} X^\sigma + (\partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho}) X^\rho + (\Gamma^\sigma_{\mu\rho} \nabla_\nu - \Gamma^\sigma_{\nu\rho} \nabla_\mu) X^\rho, \quad (0.3.93)$$

$$= (\partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho}) X^\rho + (\Gamma^\sigma_{\mu\rho} \partial_\nu - \Gamma^\sigma_{\nu\rho} \partial_\mu) X^\rho + (\Gamma^\sigma_{\mu\rho} \Gamma^\rho_{\nu\lambda} - \Gamma^\sigma_{\nu\rho} \Gamma^\rho_{\mu\lambda}) X^\lambda, \quad (0.3.94)$$

$$= X^\rho (\partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho}) + (\Gamma^\sigma_{\mu\rho} \Gamma^\rho_{\nu\lambda} - \Gamma^\sigma_{\nu\rho} \Gamma^\rho_{\mu\lambda}) X^\lambda, \quad (0.3.95)$$

where one should note that whenever a term appears after a derivative here it is to be differentiated, even if it is outside a bracket. We can introduce the Riemann tensor here from

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) X^\sigma = R^\sigma_{\rho\mu\nu} X^\rho, \quad (0.3.96)$$

and setting $X = e_\rho$, the coordinate basis vector associated with the x^ρ coordinate, the Riemann tensor can be written as

$$R^\sigma_{\rho\mu\nu} = \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho} + \Gamma^\sigma_{\mu\lambda} \Gamma^\lambda_{\nu\rho} - \Gamma^\sigma_{\nu\lambda} \Gamma^\lambda_{\mu\rho}. \quad (0.3.97)$$

Symmetries of the Riemann Tensor

Now we will discuss the symmetries of the Riemann tensor. Firstly, from the definition of the Riemann tensor, it follows that $R^\sigma_{\rho\mu\nu} = -R^\sigma_{\rho\nu\mu}$; this can be written succinctly as $R^\sigma_{\rho[\mu\nu]} = 0$. The next symmetry of the Riemann tensor will prove easy to derive using normal coordinates (described in [REF]) at a point p where $\Gamma = 0$ (but $\partial\Gamma \neq 0$) to get

$$R^\sigma_{\rho\mu\nu}|_p = \partial_\mu \Gamma^\sigma_{\nu\rho} - \partial_\nu \Gamma^\sigma_{\mu\rho}, \quad (0.3.98)$$

and it is simple to show that $R^\sigma_{[\rho\mu\nu]} = 0$ as

$$R^\sigma_{[\rho\mu\nu]}|_p = \partial_{[\mu} \Gamma^\sigma_{\nu\rho]} - \partial_{[\nu} \Gamma^\sigma_{\mu\rho]} = 0, \quad (0.3.99)$$

as the antisymmetrisation of any connection symbol like $\Gamma^\sigma_{[\nu\rho]} = 0$. Given that the tensor equation $R^\sigma_{[\rho\mu\nu]} = 0$ is true at p in normal coordinates then it is true in any coordinate system; on top of this the point p was arbitrary so therefore $R^\sigma_{[\rho\mu\nu]} = 0$ holds globally.

The next symmetry of the Riemann tensor is $R_{\sigma\rho\mu\nu} = R_{\mu\nu\sigma\rho}$. We can prove this again using normal coordinates at a point p ; here derivatives of the metric and it's inverse vanish, but second derivatives do not. The proof of the symmetry is as follows,

$$R_{\sigma\rho\mu\nu}|_p = g_{\lambda\sigma} \partial_\mu \Gamma^\lambda_{\nu\rho} - g_{\lambda\sigma} \partial_\nu \Gamma^\lambda_{\mu\rho}, \quad (0.3.100)$$

$$= \partial_\mu g_{\lambda\sigma} \Gamma^\lambda_{\nu\rho} - \partial_\nu g_{\lambda\sigma} \Gamma^\lambda_{\mu\rho}, \quad (0.3.101)$$

$$= \partial_\mu \Gamma_{\sigma\nu\rho} - \partial_\nu \Gamma_{\sigma\mu\rho}, \quad (0.3.102)$$

$$= \frac{1}{2} (\partial_\mu \partial_\rho g_{\sigma\nu} - \partial_\mu \partial_\sigma g_{\rho\nu} + \partial_\nu \partial_\sigma g_{\rho\mu} - \partial_\nu \partial_\rho g_{\sigma\mu}), \quad (0.3.103)$$

and it is a simple to show that this final expression doesn't change under swapping indices $\sigma \leftrightarrow \mu$ and $\rho \leftrightarrow \nu$.

The final symmetry of the Riemann tensor is the Bianchi identity, $\nabla_{[\lambda} R_{\sigma\rho]\mu\nu} = 0$. Using normal coordinates at a point p , we can write

$$\nabla_\lambda R_{\sigma\rho\mu\nu}|_p = \partial_\lambda R_{\sigma\rho\mu\nu}|_p \quad (0.3.104)$$

as all the christoffel symbols generated by the covariant derivative cancel and therefore

$$2\nabla_\lambda R_{\sigma\rho\mu\nu}|_p = \partial_\lambda \partial_\mu \partial_\rho g_{\sigma\nu} - \partial_\lambda \partial_\mu \partial_\sigma g_{\rho\nu} + \partial_\lambda \partial_\nu \partial_\sigma g_{\rho\mu} - \partial_\lambda \partial_\nu \partial_\rho g_{\sigma\mu}. \quad (0.3.105)$$

Antisymmetrising over λ, ρ and σ makes each term vanish as the triple partial derivatives always contain two of the antisymmetrised indices and must vanish.

To summarise, we have the following symmetries of the Riemann tensor;

$$R_{\sigma\rho[\mu\nu]} = 0, \quad (0.3.106)$$

$$R_{\sigma\rho\mu\nu} = R_{\mu\nu\sigma\rho}, \quad (0.3.107)$$

$$\nabla_{[\lambda} R_{\sigma\rho]\mu\nu} = 0. \quad (0.3.108)$$

The first two of these can be used together to give another useful relation $R_{[\sigma\rho]\mu\nu} = 0$.

Contractions of the Riemann Tensor

Now that we have explored the Riemann tensor, it is time to introduce the Ricci tensor and Ricci scalar. The Ricci tensor $R_{\mu\nu}$ is simply defined by the unique, non-zero self contraction (or trace) of the Riemann tensor,

$$R_{\rho\mu} := R^\mu_{\rho\mu\nu} = R_{\sigma\rho\mu\nu} g^{\sigma\mu}. \quad (0.3.109)$$

Contracting the the Riemann tensor with $g^{\mu\nu}$ or $g^{\sigma\rho}$ would give zero due to the antisymmetries of those indices in the tensor. Any other contractions, such as with $g^{\rho\mu}$ can be shown to be exactly the same (upto a minus sign) as contracting with $g^{\sigma\mu}$ using the symmetries of the Riemann tensor. The symmetries of the Riemann tensor guarantee that the Ricci tensor itself is symmetric. We can contract the Ricci tensor with itself (the same as taking the trace with the metric) to give us the Ricci scalar R ,

$$R = g^{\rho\nu} R_{\rho\nu}. \quad (0.3.110)$$

We can also take the trace of the Bianchi identity in Eq. (0.3.108) which gives us

$$g^{\lambda\mu} g^{\rho\nu} (\nabla_\lambda R_{\sigma\rho\mu\nu} + \nabla_\rho R_{\lambda\sigma\mu\nu} + \nabla_\sigma R_{\rho\lambda\mu\nu}) = 0, \quad (0.3.111)$$

$$\nabla^\mu R_{\sigma\mu} + \nabla^\nu R_{\sigma\nu} - \nabla_\sigma R = 0, \quad (0.3.112)$$

$$\nabla^\mu R_{\mu\sigma} - \frac{1}{2} \nabla_\sigma R = 0. \quad (0.3.113)$$

Using the contracted Bianchi identity and $\nabla \mathbf{g} = 0$ from metric compatibility, we can define the (symmetric) Einstein tensor $G_{\mu\nu}$,

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \quad (0.3.114)$$

$$\nabla^\mu G_{\mu\nu} = \nabla^\mu R_{\mu\nu} - \frac{1}{2} (g_{\mu\nu} \nabla^\mu R + \nabla^\mu (g_{\mu\nu}) R), \quad (0.3.115)$$

$$= \nabla^\mu R_{\mu\nu} - \frac{1}{2} \nabla_\nu R, \quad (0.3.116)$$

$$= 0. \quad (0.3.117)$$

The Einstein tensor $G_{\mu\nu}$ therefore has a vanishing divergence $\nabla_\mu G^{\mu\nu} = 0$.

One more useful contraction of the Riemann tensor is with a second Riemann tensor which gives the Kretschmann scalar k ,

$$k := R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}. \quad (0.3.118)$$

Both k and R are scalar fields so they are a coordinate invariant curvature measure; however section 0.4.6 will have more use of the Kretschmann scalar.

NORMAL COORDS SOMEWHERE, MAKES EXPANSIONS AND SOME DERIVATIONS EASIER, USE NORMAL COORDS FOR SYMMETRIES OF RIEMANN

MAYBE ADD THE \square AND $()$ SYMMETRY DEFINITION IN PRE-DEFINITIONS? ALSO CHECK THIS WITH THE DIFF FORMS SECTION

DEGREES OF FREEDOM

0.3.5 The Divergence Theorem

There is a generalisation of the divergence theorem to non-flat spaces, using differential geometry, which will be extremely useful in section [REF]. First, we will have to find a convenient form of the divergence $\nabla_\mu X^\mu$ of a vector \mathbf{X} . Expanding the covariant derivative gives

$$\nabla_\mu X^\mu = \partial_\mu X^\mu + \Gamma^\mu_{\mu\nu} X^\nu, \quad (0.3.119)$$

$$= \partial_\mu X^\mu + \frac{1}{2} g^{\mu\rho} (\partial_\mu g_{\rho\nu} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}) X^\nu, \quad (0.3.120)$$

$$= \partial_\mu X^\mu + \frac{1}{2} g^{\mu\rho} \partial_\nu g_{\mu\rho} X^\nu. \quad (0.3.121)$$

To simplify any further, we need to prove a matrix identity for a symmetric real matrix \mathbf{M} with determinant M ,

$$M^{-1}\partial_\mu M = M_{ij}^{-1}\partial_\mu M_{ij} = \text{Tr}(\mathbf{M}^{-1}\partial_\mu \mathbf{M}). \quad (0.3.122)$$

Simplifying the left hand side in terms of the eigenvalues λ_i of \mathbf{M} . Given that $M = \det \mathbf{M} = \prod_i \lambda_i$ we can easily write

$$M^{-1}\partial_\mu M = \partial_\mu \ln(|M|), \quad (0.3.123)$$

$$= \partial_\mu \ln \left(\left| \prod_i \lambda_i \right| \right), \quad (0.3.124)$$

$$= \partial_\mu \sum_i \ln(|\lambda_i|), \quad (0.3.125)$$

$$= \sum_i \lambda_i^{-1} \partial_\mu \lambda_i. \quad (0.3.126)$$

Now we will show that the right hand side of Eq. (0.3.122) also equals this. To do this we start by decomposing \mathbf{M} into a diagonal matrix \mathbf{D} like

$$\mathbf{M} = \mathbf{O}^{-1} \mathbf{D} \mathbf{O}, \quad (0.3.127)$$

$$\mathbf{M}^{-1} = \mathbf{O}^{-1} \mathbf{D}^{-1} \mathbf{O}, \quad (0.3.128)$$

then using the fact that $\text{Tr}(\mathbf{A}\mathbf{B}...\mathbf{C}\mathbf{D}) = \text{Tr}(\mathbf{D}\mathbf{A}\mathbf{B}...\mathbf{C})$ for matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , ... and \mathbf{D} ,

$$\text{Tr}(\mathbf{M}^{-1}\partial_\mu \mathbf{M}) = \text{Tr}(\mathbf{O}^{-1} \mathbf{D}^{-1} \mathbf{O} \partial_\mu (\mathbf{O}^{-1} \mathbf{D} \mathbf{O})), \quad (0.3.129)$$

$$= \text{Tr}(\mathbf{D}^{-1} \partial_\mu \mathbf{D}) + \text{Tr}(\mathbf{O}^{-1} \partial_\mu \mathbf{O}) + \text{Tr}(\mathbf{O} \partial_\mu \mathbf{O}^{-1}), \quad (0.3.130)$$

$$= \text{Tr}(\mathbf{D}^{-1} \partial_\mu \mathbf{D}) + \underbrace{\text{Tr}(\partial_\mu (\mathbf{O}^{-1} \mathbf{O}))}_{=0}, \quad (0.3.131)$$

$$= \text{Tr}(\mathbf{D}^{-1} \partial_\mu \mathbf{D}). \quad (0.3.132)$$

Given that \mathbf{D} is the diagonal matrix composed of the eigenvalues λ_i then it follow that,

$$\mathbf{D} = \text{Diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}, \quad (0.3.133)$$

$$\partial_\mu \mathbf{D} = \text{Diag}\{\partial_\mu \lambda_1, \partial_\mu \lambda_2, \dots, \partial_\mu \lambda_n\}, \quad (0.3.134)$$

$$\mathbf{D}^{-1} = \text{Diag}\{\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_n^{-1}\}, \quad (0.3.135)$$

then finally $\text{Tr}(\mathbf{D}^{-1} \partial_\mu \mathbf{D})$ can be evaluated in terms of the λ_i as follows,

$$\text{Tr}(\mathbf{D}^{-1} \partial_\mu \mathbf{D}) = \sum_{ij} D_{ij}^{-1} \partial_\mu D_{ij}, \quad (0.3.136)$$

$$= \sum_i D_{ii}^{-1} \partial_\mu D_{ii}, \quad (0.3.137)$$

$$= \sum_i \lambda_i^{-1} \partial_\mu \lambda_i, \quad (0.3.138)$$

which proves that Eq. (0.3.122) is true. Applying Eq. (0.3.122) to the metric \mathbf{g} gives,

$$g^{\mu\rho} \partial_\nu g_{\mu\rho} = g^{-1} \partial_\nu g, \quad (0.3.139)$$

and the covariant divergence $\nabla_\mu X^\mu$ of a vector \mathbf{X} simplifies to

$$\nabla_\mu X^\mu = \partial_\mu X^\mu + \frac{1}{2g} \partial_\mu g, \quad (0.3.140)$$

$$= \frac{1}{\sqrt{|g|}} \partial_\mu \left(\sqrt{|g|} X^\mu \right), \quad (0.3.141)$$

which in the standard pseudo-Riemannian spacetime is often written

$$\nabla \cdot \mathbf{X} = \nabla_\mu X^\mu = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} X^\mu). \quad (0.3.142)$$

This equation will have much use later in section [REF] where the divergence of a vector will be integrated over a finite n -dimensional volume M in \mathcal{M} . In order to do this, the divergence theorem of curved space is used, this is

$$\int_M \nabla \cdot \mathbf{X} \sqrt{|(n)g|} \, d^n x = \sum_i \int_{\partial M_i} \mathbf{s}^{(i)} \cdot \mathbf{X} \sqrt{|(n-1)g^{(i)}|} \, d^{n-1} x, \quad (0.3.143)$$

where the surface of M is divided into a set of $(n-1)$ -dimensional surfaces ∂M_i ; for each of these surfaces there is an $(n-1)$ -dimensional metric ${}^{(n-1)}g^{(i)}$ with determinant ${}^{(n-1)}g^{(i)}$ and unit normal vector $\mathbf{s}^{(i)}$ with ${}^{(n-1)}g^{(i)}(\mathbf{s}^{(i)}, \mathbf{s}^{(i)}) = \pm 1$. When dealing with a pseudo-Riemannian manifold, we need to take care of the direction of $\mathbf{s}^{(i)}$; in the case that $\mathbf{s}^{(i)}$ is timelike it should be in-directed, and if it's spacelike then it should be out-directed.

CHECK THE LAST STATEMENT

MAYBE DERIVE DIV THEOREM, OTHERWISE REFERENCE IT

0.4 Relativity

0.4.1 Special Relativity

In the nineteenth century, it was widely beleived that the universe was permeated by an invisible luminous aether. It was thought that light travels at a fixed speed through this aether; the aether could be thought of as a universal rest frame. A consequence of this is that moving towards/away from a light source would cause the wavelength of light to get shorter/longer in a similar fashion to the doppler effect. With that line of thought one could measure the earths speed through this rest frame by setting up an interferometer experiment. An interferometer sends a light beam through a splitter, dividing the beam into two perpendicular paths. The two beams then reach a seperate mirror and are reflected to a half-mirror that recombines the two beams after they have travelled identical distance. If the time taken for the two beams to complete their identical length journeys differs then an interference pattern will be seen at a detector placed after the half-mirror due to the phase change between the two beams. In 1887 Michelson and Morley conducted their famous interferometer experiment to measure the earth's velocity through the aether. They expected to see interference when the different beams made different angles with the velocity through the aether. However, no matter which orientation they put their experiment there was no interference pattern. This implied that light moved at exactly the same speed in any direction; this is only possible if earth is in the rest frame of the aether, but this cannot possibly be true as the planet accelerates round a circular path around a sun that is following a bigger circular path and so on. This famous result demonstrated that the speed of light was constant in all inertial frames - a result that defied newtonian mechanics. This was the first and biggest hint that a new theory of dynamics was needed.

In 1905, Einstein published "On the Electrodynamics of Moving Bodies" which contained a description of Special Relativity (SR) [CHECK THIS]. SR is essentially the idea that the laws (excluding gravity) of physics are the same in any inertial (non-accelerating) rest frame - the consequence of this is that you cannot measure the velocity of your own rest frame as no frame is special. One big problem with SR is that it does not properly describe gravity and thus describes an infinite vaccum universe; this is a problem that will be adressed in the next sections. SR alone contains many interesting results such as time dilation, length contraction and the inclusion of time into the metric.

Minkowski Space

In Newtonian physics the metric of flat space, using cartesian coordinates, is

$$g_{ij} = \delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (0.4.1)$$

with the line element

$$ds^2 = \delta_{ij} dx^i dx^j = dx^2 + dy^2 + dz^2. \quad (0.4.2)$$

In SR time is promoted to a dimension and the metric is over space and time (spacetime). The spacetime metric has a single negative eigenvalue and hence negative determinant. In cartesian coordinates, the flat spacetime metric is,

$$g_{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (0.4.3)$$

and line element

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = -c^2 dt^2 + dx^2 + dy^2 + dz^2. \quad (0.4.4)$$

where c is the speed of light which will be set to unity again. The line element is often written $ds^2 = -d\tau^2$ where τ is the *proper time* experienced by an observer. Having a metric over spacetime is a non-intuitive concept step where time is included into geometry and paths through space and time (spacetime) can have negative length. This flat spacetime is very important and is called the Minkowski spacetime.

Worldlines and Causality

Consider a particle moving through the rest frame with some world-line (path through spacetime) described by $x^\mu(\tau)$. Without loss of generality the motion can be restricted to the x axis. The world-line of this particle is given by $x^\mu = \{t, \int_{t_0}^t v(t') dt' + v(t_0), 0, 0\}$. The infinitesimal form of this description is $dx^\mu = \{dt, v dt, 0, 0\}$ and the proper time associated with the interval is

$$d\tau^2 = -\eta_{\mu\nu} dx^\mu dx^\nu = (1 - v^2) dt^2. \quad (0.4.5)$$

This equation says that if $v < 1$, where it should be remembered that the speed of light here is $c = 1$, then proper time τ and coordinate time t both flow in the same direction. Curves like this have $\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu < 0$ and are called *time-like*. If $v = 1$, which is only possible for a massless particle, then the path is called *light-like*. For $v > 1$, the path is unphysical as nothing can travel faster than the speed of light, it would correspond to an observer travelling backwards in time. A curve x^μ with $\eta_{\mu\nu} \dot{x}^\mu \dot{x}^\nu > 0$ is instead called a *space-like* curve and cannot be traversed by matter or information.

MAYBEW REWRITE THIS JUST FROM A DELTA T DELTA X PERSPECTIVE AND THEN USE IT TO DESCRIBE WORLDLINES AFTER?? MAYBE JUST DO THIS IN THE gr SECTION AS THERES NO REASON NOT TO

MAYBE DO NOW THAT WORLDLINES ARE GEODESICS?

0.4.2 Physics in Special Relativity

BOOSTS?

Many physical theories can be written using tensor calculus on this flat spacetime; the correct description of electromagnetism was itself a large reason for the development of special relativity.

The Wave Equation

One very simple example is the wave equation governing the dynamics of a scalar field ϕ ,

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \phi(x^i, t) - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \phi(x^i, t) = 0, \quad (0.4.6)$$

using cartesian coordinates and c as the speed of light which we will set to one. It should be noted that this equation only holds in one coordinate system, changing coordinate system would change the equation of motion. In SR, using the language of tensor calculus, we can write this as

$$\eta^{\mu\nu} \partial_\mu \partial_\nu \phi = 0, \quad (0.4.7)$$

which is not only much simpler but is also true in any inertial frame (using Cartesian coordinates) as is one of the postulates of special relativity.

Electromagnetism

Electromagnetism can also be cast in a very succinct form using this field theory notation. Traditionally Maxwell's equations of electromagnetism are written as

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (0.4.8)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (0.4.9)$$

$$\nabla \times \mathbf{E} + \frac{d\mathbf{B}}{dt} = 0, \quad (0.4.10)$$

$$\nabla \times \mathbf{B} - \frac{d\mathbf{E}}{dt} = \mu_0 \mathbf{J} \quad (0.4.11)$$

for charge density ρ , current density \mathbf{J} . Note these differential equations are not written using the language of differential geometry and tensor calculus. In SR, the current density and charge is promoted to a single 4-vector $j^\mu = \{\rho, j^i\}$ and the 6 degrees of freedom of the electromagnetic field are encoded in the components $F_{\mu\nu}$ of an antisymmetric tensor \mathbf{F} . The four electromagnetic potentials ϕ and A_i , defined by $E_i = -\partial_i \phi - \partial_t A_i$ and $\mathbf{B} = \nabla \times \mathbf{A}$, are also combined into one 4-vector $A_\mu = \{-\phi, A_i\}$. In cartesian coordinates, the Electromagnetic tensor \mathbf{F} is,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}, \quad (0.4.12)$$

Transforming A_μ with the electromagnetic gauge transformation like $A_\mu \rightarrow A_\mu + \partial_\mu f$, for some scalar field f , leaves the physically measureable field $F_{\mu\nu}$ unchanged,

$$F_{\mu\nu} \rightarrow \partial_\mu (A_\nu + \partial_\nu f) - \partial_\nu (A_\mu + \partial_\mu f), \quad (0.4.13)$$

$$= \partial_\mu A_\nu - \partial_\nu A_\mu + \underbrace{\partial_\mu \partial_\nu f - \partial_\nu \partial_\mu f}_{=0}, \quad (0.4.14)$$

$$= F_{\mu\nu}. \quad (0.4.15)$$

In SR the Maxwell Eqs. (0.4.8) and (0.4.11) can be identically represented as

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu, \quad (0.4.16)$$

and the other two Maxwell Eqs. (0.4.9) and (0.4.10) are identically true from computing $\partial_{[\mu} F_{\alpha\beta]}$,

$$\partial_{[\mu} F_{\alpha\beta]} = \partial_\mu \partial_\alpha A_\beta - \partial_\mu \partial_\beta A_\alpha + \partial_\alpha \partial_\beta A_\mu - \partial_\alpha \partial_\mu A_\beta - \partial_\beta \partial_\alpha A_\mu, \quad (0.4.17)$$

$$\partial_{[\mu} F_{\alpha\beta]} = 0. \quad (0.4.18)$$

talk about postulates of SR.

Boosts

stress tensors?

quick conservation law

0.4.3 Physics in Curved Space

It is a commonly known result of newtonian physics that in a rotating reference frame, with angular velocity ω , a test particle will experience three fictitious forces; the centrifugal force that grows with distance from the origin, the coriolis force that depends on the velocity of the test particle, and the Euler force depending on $\partial_t \omega$. From the point of view of an observer in the rotating frame a test particle would appear to accelerate which is a violation of Newton's first law if there is no external force. The curved path followed by the particle in the rotating reference frame is of course a constant speed straight line in the non-rotating frame, and therefore the particle follows a geodesic in the inertial frame. It is possible (but very involved) to do a coordinate transformation from the rest frame to the inertial frame and find the metric $\tilde{\eta}_{\mu\nu}$ of the rotating frame; it would then be possible to compute the geodesics in the rotating frame using Eq. (0.5.12).

In a similar way to the fictitious forces arising in rotating frames, the gravitational force can be described as a fictitious force. In the absense of rotation, the path of a particle moving without external force is seen to accelerate towards massive bodies. In other words, matter density deflects particle paths towards itself. This is a property also possessed by curved spaces; the path followed by a free particle on a non-flat spacetime can bend. MENTION MATTER IS SIMILAR TO CURVATURE THEN. The brilliant insight by Einstein was that upgrading the flat spacetime of special relativity to a curved spacetime would give a relativistic theory of gravity; this theory is general relativity. In the same way that fictitious forces arise from departing from an inertial frame with rotation, the gravitational force is a fictitious force arising from departing from an inertial frame; the difference is that an inertial frame now must follow a geodesic. This explains why an observer standing on earth experiences a reaction force, the frame of the observer is not following a geodesic, or not free falling, therefore to counteract the fictitious gravitational force generated the ground must push back with a reaction force.

Given that POSTULATE says the local laws of physics in a free falling frame are indistinguishable from special relativity, any equation of motions that we want to hold in general relativity must agree with special relativity in the low curvature limit. As a general rule, if there is a law of physics expressed as a differential equation on Minkowski space that we want to use in curved space, we must replace all partial derivatives of fields with co-variant derivatives of tensor fields; this process is called the minimal coupling approach MAYBE MAKE THIS MORE OBVIOUS WITH LAGRANGEANS?. An easy example is the wave equation from Eq. (0.4.7),

$$\eta^{\mu\nu} \partial_\mu \partial_\nu \phi = 0 \rightarrow g^{\mu\nu} \nabla_\mu \nabla_\nu \phi = 0 \quad (0.4.19)$$

where we had to replace $\partial_\mu \rightarrow \nabla_\mu$ and replace the Minkowski metric η with the curved space metric $g_{\mu\nu}$ like $\eta \rightarrow g$. Writing the laws of physics as tensor equations is extremely useful as they must hold for over all points in a curved spacetime in any coordinate system; the power of these equations is that they can be written without reference to an explicit coordinate system. If a coordinate system is then picked, assuming knowledge of $g_{\mu\nu}$, then the wave equation becomes

$$g^{\mu\nu} \nabla_\mu \nabla_\nu \phi = g^{\mu\nu} \nabla_\mu \partial_\nu \phi, \quad (0.4.20)$$

$$= g^{\mu\nu} \partial_\mu \partial_\nu \phi - g^{\mu\nu} \Gamma_{\mu\nu}^\rho \partial_\rho \phi, \quad (0.4.21)$$

in terms of partial derivatives. There is the small question of uniqueness here, what is to stop us adding arbitrary amounts and types of terms that vanish in the no curvature limit. For a simple example one is

free to choose the wave equation to be

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\phi + f(x^\mu)R^n = 0, \quad (0.4.22)$$

for constant n , function f and the Ricci scalar R discussed in REF. This equation certainly returns the regular wave equation in the low curvature limit where $R \rightarrow 0$. The general rule is to keep things simple and terms such as R which are proportional to second order derivatives of the metric are thought to be less dominant than terms such as $\Gamma^\mu_{\rho\sigma}$ which are proportional to first derivatives of the metric. Navigating this minefield of which terms to include in the laws of physics leads to the topic of modified gravity REF, the next simplest theories beyond general relativity.

As another quick example the Maxwell equations are very simple to extend to curved space

maybe principle of minimal coupling to gravity is easier to understand with lagrangeans?

do postulates of GR and SR better?

decide how to deal with geodesic / path in the early parts of this section

0.4.4 The Stress-Energy-Momentum Tensor

At the heart of field theory in physics is the stress-energy-momentum tensor \mathbf{T} , also called the energy-momentum tensor or stress tensor for short. Roughly speaking, component T^{00} is energy density, components $T^{0i} = T^{i0}$ contain energy flux or momentum density and components $T^{ij} = T^{ji}$ contain momentum fluxes. The diagonal part of T^{ij} can also be thought of as containing pressure and the off-diagonal terms containing shear stress. In flat space, the conservation of energy and momentum can be written as

$$\partial_\mu T^{\mu\nu} = 0, \quad (0.4.23)$$

in the absence of external forces. Equation (0.4.23) is also called the continuity equation and can be split into two sets of familiar equations,

$$\partial_0 T^{00} = -\partial_i T^{i0}, \quad (0.4.24)$$

$$\partial_0 T^{0j} = -\partial_i T^{ij}, \quad (0.4.25)$$

where the first equation states *"The rate of change of energy density is equal and opposite to the divergence of energy flux density"* and the second equation states *"The rate of change of momentum density is equal and opposite to the divergence of momentum flux density"*.

The stress tensor is also very useful in curved space and the continuity equation becomes,

$$\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\mu\rho} T^{\rho\nu} + \Gamma^\nu_{\mu\rho} T^{\mu\rho} = 0, \quad (0.4.26)$$

which can be rewritten using Eq. (0.3.142) as

$$\partial_\mu(\sqrt{-g}T^{\mu\nu}) = -\sqrt{-g}\Gamma^\nu_{\mu\rho}T^{\mu\rho}, \quad (0.4.27)$$

$$\partial_\mu(\mathcal{T}^{\mu\nu}) = -\Gamma^\nu_{\mu\rho}\mathcal{T}^{\mu\rho}, \quad (0.4.28)$$

where the second equation writes the stress tensor as a tensor density \mathcal{T} making the equation resemble the flat space continuity equation more closely. The take away message is that traditional continuity of energy and momentum no longer hold for curved spaces. This point will be revisited in section REF where a different type of variable is defined to try recover traditional continuity.

0.4.5 The Einstein Equation

Building on the vague notion of matter causing spacetime curvature it would be helpful to have a mathematical law saying how much curvature is caused by a matter distribution. The first guess that Einstein arrived at was to write $R_{\mu\nu} = kT_{\mu\nu}$ for some constant k . The problem is that given $\nabla_\mu T^{\mu\nu} = 0$ is the generic equation of continuity for matter, it would imply $\nabla_\mu R^{\mu\nu}$ vanishes which is not generally true. As shown in Eq. (0.3.116), the Einstein tensor $G_{\mu\nu}$ does satisfy $\nabla_\mu G^{\mu\nu} = 0$; the next simplest guess at a physical law for spacetime curvature would be $G_{\mu\nu} = kT_{\mu\nu}$. Remarkably this turns out to be correct and has successfully described all gravitational physics to date.

SOME EARLY EXPERIMENTAL STUFF LIKE PRECESSION, GRAVITATIONAL TIME DILATION, LIGHT RAY DEFLECTION, RED SHIFT FOR ATOMIC SPECTRA.

At the core of General Relativity is the Einstein Equation

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \quad (0.4.29)$$

also called the Einstein Field Equations. This equation relates spacetime curvature, encoded in the Einstein tensor $G_{\mu\nu}$, to the matter distribution described by the stress tensor $T_{\mu\nu}$. To be able to solve this equation in the presence of matter a curvature dependant equation of motion to dictate how matter moves is required. This leads nicely to Wheeler's insightful one line summary of general relativity:

"Spacetime tells matter how to move; matter tells spacetime how to curve."

This deceptively simple equation can describe an infinite amount of vastly diverse spacetime geometries including regular flat Minkowski space, black holes, stars, planets, gravitational waves and even the entire universe; to properly describe the entire universe a small modification has to be made to this equation as shown in REF.

General Relativity in Vacuum

In the case of a vacuum spacetime with $T_{\mu\nu} = 0$, or one where the matter distribution is supposed to be so small that it does not cause a spacetime curvature backreaction, the Einstein equation simplifies greatly to

$$G_{\mu\nu} = 0. \quad (0.4.30)$$

Taking the trace of the above equation and the definition of $G_{\mu\nu}$ in Eq. (REF), we can see that

$$g^{\mu\nu} G_{\mu\nu} = 0 = g^{\mu\nu} (R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}) = R(1 - \frac{D}{2}), \quad (0.4.31)$$

where D is the number of spacetime dimensions. Clearly for $D \neq 2$ a vanishing einstein implies a vanishing Ricci scalar R ; if the Ricci scalar vanishes then $G_{\mu\nu} = R_{\mu\nu}$ and the Einstein equation in vacuum simplifies to

$$R_{\mu\nu} = 0. \quad (0.4.32)$$

In the special case of $D = 2$ it can be shown that $R_{\mu\nu} = Rg_{\mu\nu}/2$ [MAYBE REF EARLIER SECTION ON RIEMANN TENSOR] and the Einstein tensor $G_{\mu\nu} \equiv 0$ which renders Einstein's equation nonsensical.

0.4.6 Black Holes

As it turns out, vacuum General Relativity can describe more than just Minkowski space. Arguably the most important family of solutions to Einstein's equations in vacuum are black holes. The first black hole solution was found by Karl Schwarzschild, who's surname fittingly means "*black shield*" in German. This solution, known as the Schwarzschild solution, was discovered in 1916 with the intention of computing

the spacetime curvature in the vacuum about a spherically symmetric mass such as stars and planets. The solution for $g_{\mu\nu}$ was assumed to take the following ansatz,

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -A(r)dt^2 + B(r)dr^2 + r^2(d\theta + \sin^2(\theta)d\phi^2) \quad (0.4.33)$$

which is manifestly spherically symmetric and static; in the case $A(r) = B(r) = 1$ the solution is exactly Minkowski space in spherical polar coordinates. Schwarzschild sought a solution of this form that at large radius as $r \rightarrow \infty$ the functions $A \rightarrow 1$ and $B \rightarrow 1$, this means at large radius the solution tends to flat space. By solving $R_{\mu\nu} = 0$ from REF, Schwarzschild found a solution of the form

$$g_{\mu\nu}dx^\mu dx^\nu = -\left(1 - \frac{2Gm}{rc^2}\right)dt^2 + \left(1 - \frac{2Gm}{rc^2}\right)^{-1}dr^2 + r^2(d\theta + \sin^2(\theta)d\phi^2), \quad (0.4.34)$$

known as the Schwarzschild solution. The constants G and c are included for completeness, but are equal to one in Planck units. The Schwarzschild solution describes the spacetime about a non-spinning sphere of mass m and as $r \rightarrow \infty$ or $m \rightarrow 0$ we do approach the vacuum Minkowski spacetime as desired.

Polar-Areal Coordinates

The type of coordinates used in the Schwarzschild solution are called polar-areal coordinates; these are coordinates that satisfy $g_{\theta\theta} = r^2$, $g_{\phi\phi} = \sin^2(\theta)r^2$ and otherwise $g_{\mu\phi} = g_{\mu\theta} = 0$ CHECK THIS. Polar areal coordinates return a length of $2\pi r_0$ when integrating the length of a complete circle at fixed radius r_0 . This can be calculated using Eq. (0.2.56) and integrating with respect to ϕ around a circle while keeping $t = t_0$, $r = r_0$ and $\theta = \pi/2$ constant,

$$\int_0^{2\pi} \sqrt{g_{\phi\phi}} \Big|_{t=t_0, r=r_0, \theta=\frac{\pi}{2}} d\phi = r_0 \sin\left(\frac{\pi}{2}\right) \int_0^{2\pi} d\phi = 2\pi r_0, \quad (0.4.35)$$

which is the same result as in regular flat space. Using Eq. (0.2.64) the surface area of a sphere with $t = t_0$, $r = r_0$ can be calculated like

$$\int_0^\pi \left[\int_0^{2\pi} \sqrt{g_{\phi\phi}g_{\theta\theta}} \Big|_{t=t_0, r=r_0} d\phi \right] d\theta = r_0^2 \int_0^\pi \left[\int_0^{2\pi} d\phi \right] \sin^2(\theta) d\theta = 4\pi r_0^2, \quad (0.4.36)$$

where $\sqrt{g_{\phi\phi}g_{\theta\theta}}$ is the determinant of the metric on the two-dimensional surface defined by $t = t_0$ and $r = r_0$. Both the circumference of a circle and the area of a sphere are the same as would be in flat space, this property comes from the metric ansatz in Eq. (0.4.33).

Coordinate Singularities and Physical Singularities

There is obviously some kind of problem at $r = 2mG/c^2$, known as the Schwarzschild radius, as g_{rr} diverges here. For any planet or star observed, the radius of the object would be much larger than the Schwarzschild metric and so the solution should be seen as invalid inside the object as it has been derived in vacuum. For a valid solution to Einstein's equations inside a spherically symmetric star or planet a non-zero stress tensor must be assumed, this is explored in more detail in REF.

What Schwarzschild didn't realise before his unfortunately early death was that his solution could in fact be trusted down to radii well inside the Schwarzschild radius $r_s = 2m$. This solution describes the eternal, non-spinning black hole of mass m called the Schwarzschild black hole. The problem at radius $r = 2m$ (where we have now set $c = G = 1$) is due to the choice of coordinates and is not physically problematic. An easy way to show this is to compute a curvature scalar for the spacetime and show that it is smooth at $r = 2m$. The first choice of curvature scalar would be the Ricci scalar, but this was assumed to be zero in vacuum so is not useful here. Another curvature scalar is the Kretschmann scalar k , defined in Eq. (0.3.118), which does not generally vanish in vacuum. Vacuum general relativity asserts

that $R_{\mu\nu} = 0$, while that guarentees that $R = 0$ it does not guarentee that $R_{\mu\nu\rho\sigma} = 0$. Calculating the Kretschmann scalar for the Schwarzschild metric gives

$$k_{sc} = \frac{48m^2}{r^6}. \quad (0.4.37)$$

As can be seen, k is continuous and infinitely differentiable at $r = 2m$ but as $r \rightarrow 0$, $k \rightarrow \infty$ and there is a real coordinate independant singularity called a physical singularity. Scalar curvature invariants are useful as if they diverge in one coordinate system then they must diverge in all coordinate systems as they do not transform under coordinate transformations. It should be noted that $\sqrt{-g}$ is a scalar density and not a true scalar so cannot be used as a scalar curvature invariant.

To remove the coordinate singularity at $r = 2m$, new coordinate singularities can be introduced. One example is to use ingoing-Eddington-Finkelstein coordinates $\{v, s, \theta, \phi\}$ defined by

$$\frac{ds}{dr} = \left(1 - \frac{2m}{r}\right), \quad (0.4.38)$$

$$v = t + s, \quad (0.4.39)$$

which transforms the line element to

$$g_{\mu\nu}x^\mu x^\nu = -\left(1 - \frac{2m}{r}\right)dv^2 + 2dvdr + r^2(d\theta + \sin^2(\theta)d\phi^2), \quad (0.4.40)$$

and as can be seen the metric no longer diverges at $r = 2m$. Being careful to notice that now the metric is not diagonal, the metric determinant can be calculated, giving $\sqrt{-g} = r^2 \sin(\theta)$. Given that the metric is finite at $r = 2m$ and the metric determinant is non-zero, the metric inverse is guarenteed to be well behaved at $r = 2m$ aswell. Therefore it has been demonstrated that the singularity at $r = 2m$ in Schwarzschild polar-areal coordinates is a coordinate singularity that vanishes when using ingoing-Eddington-Finkelstein coordinates; therefore there is no physical singularity.

Throughout this section we have ignored the fact that the inverse metric also diverges as $\theta \rightarrow 0$ or $\theta \rightarrow \pi$, this is a coordinate singularity that is present in flat space (which is equivalent to the Schwarzschild spacetime with $m = 0$). This coordinate singularity arises in flat space due to the azimuthal angle ϕ being undefined at $\theta = 0$ and $\theta = \pi$. There are no physical singularities in flat space, as you might expect, and this coordinate singularity vanishes when using cartesian coordinates.

Isotropic Coordinates

A coordinate system that will be very useful later on in REF is the isotropic coordinate system. Isotropic coordinates are have the line element,

$$g_{\mu\nu}dx^\mu dx^\nu = -\Omega(r)^2 dt^2 + \Psi^2(r) \underbrace{\left(dr^2 + r^2(d\theta^2 + \sin^2(\theta)d\phi^2)\right)}_{\text{flat space}} \quad (0.4.41)$$

where the flat space line element is multiplied by a scale factor $\Psi^2(r)$. The reason these coordinates are called isotropic coordinates is that each spatial direction is equivalent when using Cartesian coordinates, as can be seen by writing the line element,

$$g_{\mu\nu}dx^\mu dx^\nu = -\Omega(r)^2 dt^2 + \Psi^2(r) (dx^2 + dy^2 + dz^2), \quad (0.4.42)$$

where $r^2 = x^2 + y^2 + z^2$. The Schwarzschild black hole solution can be expressed in isotropic coordinates aswell,

$$g_{\mu\nu}dx^\mu dx^\nu = - \left(\frac{1 - \frac{m}{2r}}{1 + \frac{m}{2r}} \right)^2 dt^2 + \left(1 + \frac{m}{2r} \right)^4 \left(dr^2 + r^2 (d\theta^2 + \sin^2(\theta)d\phi^2) \right) \quad (0.4.43)$$

$$= - \left(\frac{1 - \frac{m}{2r}}{1 + \frac{m}{2r}} \right)^2 dt^2 + \left(1 + \frac{m}{2r} \right)^4 ds_E^2, \quad (0.4.44)$$

where ds_E^2 is the three-dimensional Euclidean space. As $r \rightarrow \infty$ the line element reduces to the Minkowski space one. The radius $r = m/2$ has the same coordinate singularity that was seen in polar areal coordinates and at first glance one might think that there is a physical singularity at $r = 0$. However, if a new radial coordiante ξ is used, where $r = \frac{m^2}{4\xi}$ for some constant, the line element becomes

$$g_{\mu\nu}dx^\mu dx^\nu = - \left(\frac{1 - \frac{m}{2\xi}}{1 + \frac{m}{2\xi}} \right)^2 dt^2 + \left(1 + \frac{m}{2\xi} \right)^4 \left(d\xi^2 + \xi^2 (d\theta^2 + \sin^2(\theta)d\phi^2) \right). \quad (0.4.45)$$

This is a remarkeable result, inverting the radial coordinate about $r = 2/m$ has returned an exactly identical metric. Given that at $r = \infty$ we have flat space this implies that at $\xi = \infty$ (or $r = 0$) there is another separate flat space, not a physical singularity might have been thought. This second flat space as $\xi \rightarrow \infty$ is often called the alternate (PARALLEL MAYBE) universe and they are joined by the intraversable Einstein-Rosen bridge at $r = 2/m$. The reason that the physical singularity at $r = 0$ does not appear in isotropic coordinates is desceptively simple; in isotropic coordinates $r = 0$ does not correspond to the same point on the manifold as $r = 0$ does in the polar areal gauge for any value of t , θ or ϕ . Infact, the physical singularity of the black hole is outside of the patch on the manifold covered by isotropic coordinates. The idea of pathes is covered later EXPAND THIS WITH PENROSE MAYBE/

As described in section REF, spherically symmetric non-vacuum solutions to Einstein's equations can be found using isotropic coordinates; these solutions represent stars, planets and exotic balls of matter. The isotropic gauge is used when superposing these solutions later on in sections REF. EXPAND THIS

Penrose Diagrams and the Causal Structure of the Schwarzschild Black Hole

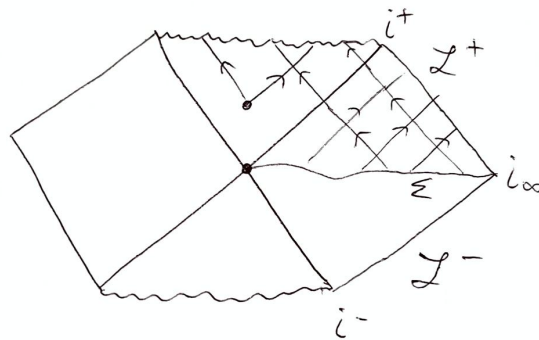


Figure 2: Diagram for proof of Penrose diagram of a Schwarzschild diagram.

Spinning, Charged Black Holes and Higher Dimensions

do polar areal, then transform to isotropic for use later. then talk about other types of black holes. talk about the horizon maybe and causal disconnect

The subject of black holes in General Relativity is huge.

reference kerr don't bother writing it

grav time dilation - maybe put this in the general GR bit, along with precession mercuries orbit?

ref higher dimensions maybe

BIRKOFFS THEOREM

NEED A PROPER BH SOLUTION SECTION IN THE NR SECTION THAT HAS ALL THE NUMERICAL APPROACHES

0.4.7 The Cosmological Constant

A discussion of general relativity is incomplete without discussing the cosmological constant Λ . At a geometric level, the cosmological constant encodes the homogeneous spacetime curvature in the absence of matter, and indeed setting $\Lambda \rightarrow 0$ (in vacuum) returns asymptotically flat vacuum general relativity. The cosmological constant is added into Einstein's equation, Eq. (0.4.29), with a term like $\Lambda g_{\mu\nu}$,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}. \quad (0.4.46)$$

Each term still has a zero-divergence as $\nabla_\mu g_{\alpha\beta} = 0$ due to the Levi-Civita connection defined in section 0.3.3. The Einstein equation in vacuum becomes

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 0, \quad (0.4.47)$$

and the trace becomes

$$R = \frac{2D}{D-2}\Lambda \quad (0.4.48)$$

for D spacetime dimensions. This quite nicely shows us that in the limit $\Lambda = 0$ the normal vacuum GR solution $R = 0$ is returned. In the case $\Lambda \neq 0$ it describes a homogeneous curved universe with constant non-zero R .

Friedmann–Lemaître–Robertson–Walker Spacetime

To model the entire universe more realistically, an isotropic uniform fluid is added. A uniform matter distribution is often done with a perfect fluid with stress tensor

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu + Pg_{\mu\nu}, \quad (0.4.49)$$

where ρ and P are the rest frame density and pressure. The fluid is taken to be at rest with $u^i = 0$ and $g(\mathbf{u}, \mathbf{u}) = -1$. The line element takes the following ansatz

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j, \quad (0.4.50)$$

where $a(t)$ is the scale factor of the universe and the spacelike metric γ is given by

$$\gamma_{ij}dx^i dx^j = \begin{cases} dr^2 + \sin^2(r) (d\theta^2 + \sin^2(\theta)d\phi^2) & , \text{ (closed)} \\ dr^2 + r^2 (d\theta^2 + \sin^2(\theta)d\phi^2) & , \text{ (flat)} \\ dr^2 + \sinh^2(r) (d\theta^2 + \sin^2(\theta)d\phi^2) & , \text{ (hyperbolic)} \end{cases} \quad (0.4.51)$$

for either uniformly curved hyperbolic space, uniformly curved closed space or flat space. This spacetime is called the Friedmann–Lemaître–Robertson–Walker spacetime and the trace of Einstein’s equation in four dimensions becomes,

$$R = 4\Lambda - 8\pi T, \quad (0.4.52)$$

where planck units are used and $T = T_{\mu\nu}g^{\mu\nu}$. It should be noted that R and T are constant over the entire spatial domain of the spacetime for a moment of time. This form of Einstein’s equation makes it abundantly clear that the cosmological constant has the same effect on curvature as a uniform matter distribution. In the case $4\Lambda = 8\pi T$, Einstein’s equation is once again $R = 0$ and can be solved with the flat metric of Minkowski (but now it is not in vacuum unless $T = \Lambda = 0$). If $8\pi T > 4\Lambda$, then the spatial hypersurface of the universe is closed and finite sized. Finally, if $4\Lambda > 8\pi T$ then the spatial hypersurface of the universe is open or hyperbolic; as a radial geodesics is followed, the amount of space grows faster than the r^2 as expected in flat space.

Putting the ansatz for the metric and stress tensor into the Einstein equation returns an ODE for $a(t)$. Along with an equation of state, $P(\rho)$, the continuity equation $\nabla_\mu T^{\mu\nu} = 0$ returns an ODE for $\rho(t)$. The two solutions $\rho(t)$ and $a(t)$, along with knowledge of whether the spatial time-slices are flat, closed or hyperbolic, fully specify the spacetime. These solutions cover many topics such as the big bang, inflation, universe expansion, bouncing cosmologies (ending with a big crunch) and also minkowski space given the correct conditions.

0.4.8 The Lagrangean Formulation of General Relativity

A common procedure in theoretical physics is to encapsulate the solution space in an action functional S ,

$$S = \int \mathcal{L} \sqrt{-g} \, dx^4. \quad (0.4.53)$$

Using the calculus of variation on this action returns differential equations governing the system. As found by Hilbert [REF] the following lagrangean density $\mathcal{L} = R$, equal to the Ricci scalar, returns the vacuum Einstein equations under varying with respect to $g^{\mu\nu}$.

$$\delta S = \int [\sqrt{-g}(\delta R) + R(\delta\sqrt{-g})] \, dx^4 \quad (0.4.54)$$

$$= \int \left[\sqrt{-g} \delta(g^{\mu\nu} R_{\mu\nu}) - \frac{1}{2} \sqrt{-g} g_{\mu\nu} R(\delta g^{\mu\nu}) \right] \, dx^4 \quad (0.4.55)$$

$$= \int \left[\sqrt{-g} g^{\mu\nu} (\delta R_{\mu\nu}) + \sqrt{-g} \left(R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) \delta g^{\mu\nu} \right] \, dx^4 \quad (0.4.56)$$

where we used Eq. (0.3.139) to vary $\sqrt{-g}$. Remembering that the difference of two christoffel symbols such as,

$$\delta \Gamma_{\mu\nu}^\lambda = \Gamma_{\mu\nu}^\lambda|_{g^{\mu\nu} + \delta g^{\mu\nu}} - \Gamma_{\mu\nu}^\lambda|_{g^{\mu\nu}} \quad (0.4.57)$$

is a tensor, and using normal coordinates [WRITE UP NORMAL COORDS] the left hand term of Eq. (0.4.56) becomes

$$g^{\mu\nu} \delta R_{\mu\nu} = g^{\mu\nu} \left(\partial_\lambda \delta \Gamma_{\mu\nu}^\lambda - \partial_\nu \delta \Gamma_{\mu\lambda}^\lambda \right), \quad (0.4.58)$$

$$= g^{\mu\nu} \left(\nabla_\lambda \delta \Gamma_{\mu\nu}^\lambda - \nabla_\nu \delta \Gamma_{\mu\lambda}^\lambda \right), \quad (0.4.59)$$

$$= \nabla_\lambda \left(g^{\mu\nu} \delta \Gamma_{\mu\nu}^\lambda - g^{\mu\lambda} \delta \Gamma_{\mu\nu}^\nu \right), \quad (0.4.60)$$

$$= \nabla_\lambda X^\lambda. \quad (0.4.61)$$

The ability to transform the partial derivatives into covariant derivatives comes from using normal coordinates. Putting everything together, δS becomes

$$\delta S = \int \left[\nabla_\mu X^\mu + \left(R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) \right] \sqrt{-g} dx^4, \quad (0.4.62)$$

$$= \int_B X^\mu \hat{s}_\mu \sqrt{|^{(3)}g|} dx^3 + \int \left[R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right] \sqrt{-g} dx^4, \quad (0.4.63)$$

$$(0.4.64)$$

where the integral over B represents the surface integral over the boundary of our spacetime with metric $^{(3)}g_{ij}$; on B $\delta g^{\mu\nu} \rightarrow 0$ and therefore $X^\mu \rightarrow 0$. Setting $\delta S = 0$ implies

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0, \quad (0.4.65)$$

which is the vacuum Einstein equation.

Non-Vacuum Spacetimes

Matter is often added into a spacetime at the level of the lagrangean with a term $\frac{16\pi G}{c^4} \mathcal{L}_m$. The cosmological constant can be added in exactly the same way with \mathcal{L}_Λ . The total lagrangean becomes,

$$S = \int \left(R + \frac{16\pi G}{c^4} \mathcal{L}_m + \mathcal{L}_\Lambda \right) \sqrt{-g} dx^4. \quad (0.4.66)$$

As we have already seen earlier in this section, the variation of R with respect to the inverse metric components $g^{\mu\nu}$ returns the vacuum Einstein equation; adding the two new terms from \mathcal{L}_m and \mathcal{L}_Λ , setting $\delta S = 0$ gives different equation

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \frac{16\pi G}{c^4} \frac{1}{\sqrt{-g}} \frac{\delta(\mathcal{L}_m \sqrt{-g})}{\delta g^{\mu\nu}} + \frac{1}{\sqrt{-g}} \frac{\delta(\mathcal{L}_\Lambda \sqrt{-g})}{\delta g^{\mu\nu}} = 0. \quad (0.4.67)$$

Comparing the \mathcal{L}_Λ term to the Einstein equation with cosmological constant in Eq. (0.4.46) we must have

$$g_{\mu\nu} \Lambda = \frac{1}{\sqrt{-g}} \frac{\delta(\mathcal{L}_\Lambda \sqrt{-g})}{\delta g^{\mu\nu}}, \quad (0.4.68)$$

$$= \mathcal{L}_\Lambda \frac{1}{\sqrt{-g}} \frac{\delta(\sqrt{-g})}{\delta g^{\mu\nu}} + \frac{\delta(\mathcal{L}_\Lambda)}{\delta g^{\mu\nu}}, \quad (0.4.69)$$

$$= -\frac{1}{2} g_{\mu\nu} \mathcal{L}_\Lambda + \frac{\delta(\mathcal{L}_\Lambda)}{\delta g^{\mu\nu}}, \quad (0.4.70)$$

$$(0.4.71)$$

which is solved by $\mathcal{L}_\Lambda = -2\Lambda$. Comparing the matter term instead returns another definition of the stress tensor,

$$T_{\mu\nu} := \frac{2}{\sqrt{-g}} \frac{\delta(\mathcal{L}_m \sqrt{-g})}{\delta g^{\mu\nu}}. \quad (0.4.72)$$

Collecting these results, the full lagrangean is,

$$\mathcal{L} = R + \frac{16\pi G}{c^4} \mathcal{L}_m - 2\Lambda, \quad (0.4.73)$$

or in Planck units,

$$\mathcal{L} = R + 16\pi \mathcal{L}_m - 2\Lambda. \quad (0.4.74)$$

The form of \mathcal{L}_m is problem specific, depending on the type of matter. The equation of motion of the matter, described by a set of fields ϕ_i , and their partial derivatives with respect to x^μ is,

$$\frac{\delta \mathcal{L}}{\delta \phi_i} - \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi_i} \right) = 0. \quad (0.4.75)$$

If the ϕ_i are scalar fields then the equation of motion simplifies to

$$\frac{\delta \mathcal{L}}{\delta \phi_i} - \nabla_\mu \left(\frac{\delta \mathcal{L}}{\delta \nabla_\mu \phi_i} \right) = 0. \quad (0.4.76)$$

Modified Theories of Gravity

CAN HELP WITH MORE THEORETICAL THEORIES, NEXT EXPANSIONS OF R SUCH AS MODIFIED GRAV OR CHERNS SIEMANS OR HORNDENSKI.

[ADD SOME MOTIVATION OF WHY WE LIKE LAGRANGIANS, EASY TO QUANTIZE OR ADD MORE MATTER OR INTERACTIONS OF MATTER? note hilbert did einstein eq before einstein]

0.5 Stuff

0.5.1 notes

prove the diff of christoffels is a tensor?

NEWman penrose formalism?

wheeler quote, matter tells space how to curve, and space tells matter how to move.

GRChombo section?

einstein summation conventions? and upstairs/downstairs indices in conventions

distinguish between tensors and tensor fields (especially in the d/dlambda and X bit)

maybe work coordinate transformation into the text earlier with the introduction of vectors? maybe split the long diff geom section (functions curves and tensors) into two sections, functions curves, coords trans, vectors part 1 then covectors, tensors and tensor transformation part 2, forms part 3 then differentiation then calculus?

mention SR and the fact that GR is based on the equivalence principles (weak and strong?) somewhere.

maybe make SR as a solution to the Einstein equation a different section to just SR.

decide when to use SR and GR acronyms

1887 Michelson-Morley to find aether that light moves on. Instead found not true. laws of physics are the same in all frames. Assuming these two things Einstein derived the Lorentz group.

connection/christoffel and capitalise or not?

talk about newman penrose scalars for GW extraction?

check capital letters

define trace

3+1 stress tensor is wrong for klein gordon

adm metric is wrong in 4th term report

add a sections with intro reading such as books alcubbiere baumgarte, adn GR stuff like sean caroll, tong, harvey ..

use invisible (with star) subsections to break up longer sections

0.5.2 Differential Forms

A differential p -form is an antisymmetric $(0, p)$ tensor; the tensor components (e.g. $A_{\alpha\beta\dots\zeta}$) vanish if there is a repeating index, equal 1 for an even permutation of indeces (like 1,2,3,...,N) and -1 for an odd permutation (such as 2,1,3,4,...,N). Note that a 0-form is a function, a 1-form is a co-vector and for an N -dimensional manifold the N -form is unique and the p -forms with $p > N$ vanish.

When considering differential forms, the conventional covector basis is often changed from θ^μ to dx^μ which will lend itself nicely to integrating p-forms on manifolds later. This convention means we can write the metric as $\mathbf{g} = g_{\mu\nu}dx^\mu \otimes dx^\nu$; the same can be done for any tensor. Note that the metric cannot be a 2-form as it is not a symmetric tensor, infact the metric is a symmetric tensor.

Differential forms have their own unique differential operator called the exterior derivative, it acts on a p -form and returns a $p + 1$ -form. For a p -form $\mathbf{A} = A_\mu dx^\mu$, the exterior derivative is given by

$$(dA)_{\mu_1\dots\mu_{p+1}} = (p+1)\partial_{[\mu_1}A_{\mu_2\dots\mu_{p+1}]}, \quad (0.5.1)$$

where the square brackets mean the antisymmetric [check as i might need epsilons here]. This derivative guarentees to return a tensor without the need for a connection or metric on the manifold, unlike the partial derivative that we will see next. One common example of an exterior derivative is of a 1-form, say \mathbf{A} , giving

$$(dA)_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (0.5.2)$$

which returns a 2-form. To be a 2-form the tensor must be antisymmetric under swapping indeces like $(dA)_{\mu\nu} = -(dA)_{\nu\mu}$, which is trivially true, and must transform like a tensor. Performing a coordinate transformation on $(dA)_{\mu\nu}$ we see it transforms like

$$\frac{\partial}{\partial \tilde{x}^\mu} \tilde{A}_\nu - \frac{\partial}{\partial \tilde{x}^\nu} \tilde{A}_\mu = \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\rho} \left(\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} A_\sigma \right) - \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \frac{\partial}{\partial x^\sigma} \left(\frac{\partial x^\rho}{\partial \tilde{x}^\mu} A_\rho \right), \quad (0.5.3)$$

$$= \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \left(\frac{\partial}{\partial x^\rho} A_\sigma - \frac{\partial}{\partial x^\sigma} A_\rho \right) + \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial}{\partial x^\rho} \left(\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \right) A_\sigma - \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \frac{\partial}{\partial x^\sigma} \left(\frac{\partial x^\rho}{\partial \tilde{x}^\mu} \right) A_\rho, \quad (0.5.4)$$

$$= \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \left(\frac{\partial}{\partial x^\rho} A_\sigma - \frac{\partial}{\partial x^\sigma} A_\rho \right) + \left(\frac{\partial^2 x^\rho}{\partial \tilde{x}^\mu \partial \tilde{x}^\nu} - \frac{\partial^2 x^\rho}{\partial \tilde{x}^\nu \partial \tilde{x}^\mu} \right) A_\rho, \quad (0.5.5)$$

$$= \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} \left(\frac{\partial}{\partial x^\rho} A_\sigma - \frac{\partial}{\partial x^\sigma} A_\rho \right), \quad (0.5.6)$$

which is exactly the tensor transformation law.

COVER HER THAT DD GIVES ZERO? HODGE ? REWRITE DIFF EQNS WITH FORM NOTATION AND HODGE THEORY? PROVE

0.5.3 Parallel Transport

Parallel transport is the procedure of moving tensors along a smooth curve in a manifold. The curve Γ can be expressed parametrically as $x^\mu(\tau)$ with respect to a parameter τ , this automatically gives us the tangent vector \mathbf{X} to Γ like

$$X^\mu(\tau) = \frac{\partial x^\mu(\tau)}{\partial \tau} = \dot{x}^\mu. \quad (0.5.7)$$

Parallel transporting a tensor \mathbf{T} along Γ requires that $\nabla_{\mathbf{X}}\mathbf{T}$ must be satisfied along Γ where $\nabla_{\mathbf{X}} = x^\mu \nabla_\mu$.

MAYBE DO THE FAMOUS VECTOR ON A SPHERE EXAMPLE

Having discussed parallel transport, we can re-derive geodesics. Take a vector field \mathbf{X} and parallel transport is along it's own integral curves, therefore \mathbf{X} must satisfy $\nabla_{\mathbf{X}} \mathbf{X} = 0$. With a little algebra we can show that this describes a geodesic,

$$x^\mu \nabla_\mu X^\nu = 0, \quad (0.5.8)$$

$$x^\mu \nabla_\mu X^\nu = X^\mu \partial_\mu X^\nu + X^\mu \Gamma^\nu_{\rho\mu} X^\rho, \quad (0.5.9)$$

$$= \frac{\partial x^\mu(\tau)}{\partial \tau} \frac{\partial}{\partial x^\mu} \frac{\partial x^\nu(\tau)}{\partial \tau} + \Gamma^\nu_{\rho\mu} \frac{\partial x^\mu(\tau)}{\partial \tau} \frac{\partial x^\rho(\tau)}{\partial \tau}, \quad (0.5.10)$$

$$= \frac{\partial}{\partial \tau} \frac{\partial x^\nu(\tau)}{\partial \tau} + \Gamma^\nu_{\rho\mu} \frac{\partial x^\mu(\tau)}{\partial \tau} \frac{\partial x^\rho(\tau)}{\partial \tau}, \quad (0.5.11)$$

$$= \ddot{x}^\nu + \Gamma^\nu_{\rho\mu} \dot{x}^\rho \dot{x}^\mu \quad (0.5.12)$$

and we have re-derived Eq. (0.2.71). Again this aligns with our intuition of a geodesic in flat space. We already remarked in Section 0.2.6 that $\ddot{x}^\mu = 0$ describes a straight line in flat space (when using cartesian coordinates) and now we can add another interpretation; in flat space a geodesic (which is a straight line) can be created by transporting a vector along it's own direction.

0.5.4 Low Curvature Limit of General Relativity

An important use of General Relativity is it's use in the low curvature limit. The simplest example of this would be Special Relativity; if General Relativity [DO] reproduce Special Relativity in the limit of vanishing curvature then it is. We work with the assumption of a vacuum spacetime with no Cosmological constant and seek solutions to the Einstein equation with metric $g_{\mu\nu} = \eta_{\mu\nu}$ where

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (0.5.13)$$

in Cartesian coordinates. As seen before, in Eq.[REF]REF the Einstein equation in vacuum simplfies to just a vanishing Ricci tensor,

$$R_{\mu\nu} = 0, \quad (0.5.14)$$

$$= \partial_\rho \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\mu\rho} + \Gamma^\rho_{\rho\sigma} \Gamma^\sigma_{\mu\nu} - \Gamma^\rho_{\mu\sigma} \Gamma^\sigma_{\rho\nu}. \quad (0.5.15)$$

Given that the Connection symbols $\Gamma^\mu_{\nu\rho}$ vanish everywhere for the metric components $\eta_{\mu\nu}$ then Eq. (0.5.14) is trivially satisfied and we have proved that Special Relativity is the zero-curvature limit of GR.

Relaxing the condition $g_{\mu\nu} = \eta_{\mu\nu}$ to $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where the components $h_{\mu\nu} \ll 1$, we can create a vacuum spacetime consisting of small curvature fluctuations; this turns out to describe gravitational waves. Ignoring terms of order $\mathcal{O}(h^2)$ we can see that

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}, \quad (0.5.16)$$

$$h^{\rho\sigma} = \eta^{\mu\rho} \eta^{\nu\sigma} h_{\mu\nu}, \quad (0.5.17)$$

$$g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho = \eta^{\mu\nu} \eta_{\nu\rho} + \eta^{\mu\nu} h_{\nu\rho} - h^{\mu\nu} \eta_{\nu\rho} + \mathcal{O}(h^2), \quad (0.5.18)$$

$$= \delta^\mu_\rho + \eta^{\mu\nu} h_{\nu\rho} - \eta^{\mu\alpha} \eta^{\nu\beta} h_{\alpha\beta} \eta_{\nu\rho}, \quad (0.5.19)$$

$$= \delta^\mu_\rho + \underbrace{\eta^{\mu\nu} h_{\nu\rho} - \eta^{\mu\alpha} h_{\alpha\rho}}_{=0}. \quad (0.5.20)$$

Note that we raise/lower the indices of $h_{\mu\nu}/h^{\mu\nu}$ with η and not g THIS IS WRONG. Looking at the connection symbols and Ricci tensor while ignoring $\mathcal{O}(h^2)$ terms we get

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}(\eta^{\rho\sigma} - h^{\rho\sigma})(\partial_\mu(\eta_{\sigma\nu} + h_{\sigma\nu}) + \partial_\nu(\eta_{\mu\sigma} + h_{\mu\sigma}) - \partial_\sigma(\eta_{\mu\nu} + h_{\mu\nu})), \quad (0.5.21)$$

$$= \frac{1}{2}\eta^{\rho\sigma}(\partial_\mu h_{\sigma\nu} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}) + \mathcal{O}(h^2), \quad (0.5.22)$$

$$R_{\mu\nu} = \partial_\rho \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\mu\rho} + \mathcal{O}(h^2), \quad (0.5.23)$$

$$= \frac{1}{2}\eta^{\rho\sigma}(\partial_\rho \partial_\mu h_{\sigma\nu} - \partial_\rho \partial_\sigma h_{\mu\nu} - \partial_\nu \partial_\mu h_{\sigma\rho} + \partial_\sigma \partial_\mu h_{\rho\mu}) + \mathcal{O}(h^2), \quad (0.5.24)$$

$$R = (\eta^{\mu\nu} - h^{\mu\nu})R_{\mu\nu}, \quad (0.5.25)$$

$$= \eta^{\mu\nu}\eta^{\rho\sigma}(\partial_\mu \partial_\rho h_{\sigma\nu} - \partial_\sigma \partial_\rho h_{\mu\nu}) + \mathcal{O}(h^2). \quad (0.5.26)$$

To simplify the solving of Einsteins eqn [FIX THIS] we can consider coordinate transformation like $x^\mu \rightarrow x^\mu + \zeta^\mu$ where $\zeta^\mu \ll 1$. It can easily be shown that to first order in ζ^μ and $h_{\mu\nu}$ that $h_{\mu\nu}$ transforms like

$$h_{\mu\nu} \rightarrow h_{\mu\nu} + \partial_\mu \zeta_\nu + \partial_\nu \zeta_\mu \quad (0.5.27)$$

where raising/lowering the indices of ζ is done with η as terms of order $\mathcal{O}(h)\mathcal{O}(\zeta)$ are ignored.

Note [TALK ABOUT G PLUS H RATHER THAN ETA PLUS H FOR BH RINGDOWN AND MAYBE BS STABILITY? THIS IS MORE PERTURBATION THEORY THOUGH]

[MENTION : GW'S, PRECESSION, LIGHT DEFLECTION, TIME DILATION AROUND EARTH/-SUN, MAYBE MENTION INTERSTELLAR'S TIME DILATION? ARE THESE REALLY LOW ENERGY LIMIT? MENTION POST NEWTONIAN OTHER THINGS?]