Phys 476

GENERAL RELATIVITY

University of Waterloo

Course notes by: TC Fraser
Instructor: Florian Girelli

tcfraser@tcfraser.com

Version: 1.2 - Course Complete

Table Of Contents

Disclaimer

These notes are intended to be a reference for my future self (TC Fraser). If you the reader find these notes useful in any capacity, please feel free to use these notes as you wish, free of charge. However, I do not guarantee their complete accuracy and mistakes are likely present. If you notice any errors please email me at **tcfraser@tcfraser.com**, or contribute directly at **https://github.com/tcfraser/course-notes**. If you are the professor of this course (Florian Girelli) and you've managed to stumble upon these notes and would like to make large changes or additions, email me please.

Latest versions of all my course notes are available at **www.tcfraser.com/coursenotes**.

1 Introduction

1.1 History

The first lecture was a summary of astrophysical history from around ∼200BC to today. I elected not to take notes as it was pretty standard stuff and a lot of slides. Sorry.

2 Tensor Formalism

At the core of General Relativity is the mathematics of differential geometry. Differential geometry requires the idea of tensors, a generalization of vectors and matrices and forms that can handle messy geometries and metrics.

Let V be a vector space of finite dimension. Any V is isomorphic to $\mathbb{R}^{n+1}$ through the coefficients of a chosen basis. Let the basis of V be given by,

$$\{e_i\}_{i=0,\ldots,n}$$

Then any vector $v \in V$ is expressible by,

$$v = \sum_{i=0}^{n} v^i e_i$$

Where v^i are the i-th coefficients of the vector v with respect to the basis $\{e_i\}$.

2.1 Einstein Summation Rule

For convenience let's provide a new, shorter notation for the vector v.

$$v^i e_i = v^0 e_0 + \ldots + v^n e_n = \sum_{i=0}^{n} v^i e_i$$

Effectively, we have just **dropped the summation sign**. The Einstein summation rule is as follows:

If there are two identical indices, 1 "up" and 1 "down", it means that a summation is secretly present, it's just be removed for convenience. Note that the i in this case is *dummy index*.

$$v^i e_i = v^\alpha e_\alpha = v^j e_j$$

Here v^i are the components of vector $v \in V$ and are real numbers. $v^i \in \mathbb{R}, \forall i \in \{0, \ldots, n\}$.

Note v^i is called the vector v when i is the set $\{0, \ldots, n\}$, but can also be called the i-th component of v when i has a fixed value $i \in \{0, \ldots, n\}$.

2.2 Examples of Basis for V

The values of e_i or the i's themselves can take on many possible values.

- Cartesian coordinates t, x, y, z
- spherical coordinates t, r, ϕ, θ
- etc.

Each of the above examples is the space $V = \mathbb{R}^4$ (with some bounds for spherical coordinates).

2.3 Dual Vector Space

The dual vector space of V denoted V^* is also isomorphic to $\mathbb{R}^{n+1}$ and is built from the space of linear forms on V.

$$V^* = \{w : V \to \mathbb{R} \mid w(\alpha v_1 + \beta v_2) = \alpha w(v_1) + \beta w(v_2)\}$$

where $v_1, v_2 \in V$ and $\alpha, \beta \in \mathbb{R}$.
In Quantum Mechanics, the vectors are the bras and the elements of the dual space (called the co-vectors) are the kets.

We note,

$$\left\{f^i\right\}_{i=0,\ldots,n}$$

is the basis for V^* is defined by the Kronecker symbol δ,

$$f^j(e_i) = \delta^j{}_i$$

$$\delta^j{}_i = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

An element in V^* is $w = w_i f^i$. w_i are the components of the covector w. Note that for a **finite dimensional vector space**,

$$V^{**} = V$$

2.4 Bilinear Maps

Introduce a bilinear map $B(v, w)$ where $B : V \times V \to \mathbb{R}$ where,

$$B(\alpha v_1 + \beta v_2, w) = \alpha B(v_1, w) + \beta B(v_2, w)$$

and the same for the other parameter w.

Examples include the inner product (otherwise known as the scalar or dot product).
Bilinear forms are bilinear maps such that the following conditions are true:

- symmetric: $B(v, w) = B(w, v)$
- non-degenerated: $B(v, w) = 0 \quad \forall v \implies w = 0$

Playing with indices,

$$\begin{aligned} B(v, w) &= B(v^\alpha e_\alpha, w^\beta e_\beta) \\ &= v^\alpha B(e_\alpha, w^\beta e_\beta) \quad \text{By linearity} \\ &= v^\alpha w^\beta B(e_\alpha, e_\beta) \quad \text{By linearity} \end{aligned}$$

A bilinear map used in this way provides a way to eliminate the headache of complicated cross sums. Define new notation,

$$B(e_\alpha, e_\beta) \equiv g_{\alpha\beta}$$

Where $g_{\alpha\beta}$ is a real number $\mathbb{R}$ because α and β are summed over.

$$B(v, w) = v^\alpha w^\beta g_{\alpha\beta} = v^\alpha g_{\alpha\beta} w^\beta = w^\beta g_{\alpha\beta} v^\alpha$$

All of the above terms are commutative because in the end, it represents a sum over all α, β.

$$B(v,w) = \underbrace{v^0 w^0 g_{00} + \ldots + v^2 w^3 g_{2,3} + \ldots + v^n w^n g_{nn}}_{(n+1)^2 \text{terms}}$$

2.5 Distance and Norms

To define a distance in a vector space, we can use norms. In this case, $g_{\alpha\beta}$ would be called the metric. The Euclidean metric (with respect to a Cartesian basis) for example would be,

$$g_{\alpha\beta} = \begin{cases} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}$$

We can also choose to enforce that the basis be orthonormal,

$$B(e_i, e_j) = \begin{cases} \pm 1 & i = j \\ 0 & i \neq j \end{cases}$$

Note that the potential for a negative norm means the notion of positive definiteness is no longer guaranteed.

2.6 Signatures of Metrics

We call the signature of the metric the number of $+1$'s and -1's appearing in g_{ij} when dealing with the orthonormal basis. Signature is denoted as:

$$(p,q) = \left(\underbrace{p}_{\text{postive}}, \underbrace{q}_{\text{negative}} \right)$$

For example,

- Euclidean metric: $(n+1, 0)$
- Minkowski metric: $(n, 1)$

Note the order of the signature is chosen to be (p,q) and not (q,p) by convention.

2.7 Co-vectors from Vectors

Note that v^i was called the vector and w_i was called the covector. This notation seems to indicate that conversion between V and V^* is notationally equivalent to raising and lowering the indices.

We call the following operation "Lowering the index using the metric".

$$\underbrace{v^\alpha}_{\text{components of vector}} \mapsto g_{\alpha\beta} v^\beta = \underbrace{v_\alpha}_{\text{components of covector}}$$

In use,

$$B(v,w) = v^\alpha g_{\alpha\beta} w^\beta = \underbrace{v_\beta}_{\text{bra}} \underbrace{w^\beta}_{\text{ket}}$$

2.8 Linear Map on V to V

$$M : V \to V$$

Where M is a matrix. An the map is equivalent to $v \to Mv \in V$. Some definition,

$$(Mv)^\alpha = \underbrace{M^\alpha{}_\beta}_{\text{Matrix(components)}} v^\beta$$

Note that $M^\alpha{}_\beta \in \mathbb{R}$ for α and β fixed. Example: The identity matrix is denoted $\delta^\alpha{}_\beta = \mathbb{I}$.

2.9 Scalar Product on Dual Space

Introduce a scalar product for the co-vectors w.

$$w, t \in V^*$$

$$w \cdot t = w_\alpha h^{\alpha\beta} t_\beta$$

Where $h^{\alpha\beta}$ is symmetric and non-degenerate.
So how is the scalar product between the dual and normal space related? Specifically how are $g_{\alpha\beta}$ and $h^{\alpha\beta}$ connected? Well,

$$\begin{aligned} v^\alpha g_{\alpha\beta} w^\beta &= v^\alpha w_\alpha \\ &= v_\gamma h^{\gamma\alpha} w_\alpha \\ &= v^\nu g_{\nu\gamma} h^{\gamma\alpha} w_\alpha \\ &= v^\nu g_{\nu\gamma} h^{\gamma\alpha} g_{\alpha\mu} w^\mu \end{aligned}$$

Since this is true for any v and w we require that,

$$h^{\gamma\alpha} g_{\alpha\mu} = \delta^\gamma{}_\mu$$

This means we say that the metric h is the inverse of the metric g. Convention on V^*: we denote the metric $g^{\alpha\beta}$ (the indices are "up").

2.10 Invariance of Scalar Product

Let us say we have a matrix $M : v \to \tilde{v} = Mv, w \to \tilde{w} = Mw$ and that M preserves the scalar product.

$$\tilde{v} \cdot \tilde{w} = v \cdot w \qquad \forall v, w$$

Examine,

$$M^\gamma{}_\alpha v^\alpha g_{\alpha\beta} M^\beta{}_\rho w^\rho = v^\alpha g_{\alpha\beta} w^\beta$$

Use commutativity and dummy-ness of indices to obtain,

$$v^\alpha M^\gamma{}_\alpha g_{\alpha\rho} M^\rho{}_\beta w^\beta = v^\alpha g_{\alpha\beta} w^\beta$$

Drop outer co-vectors v and w to get,

$$M^\gamma{}_\alpha g_{\alpha\rho} M^\rho{}_\beta = g_{\alpha\beta} \tag{2.1}$$

Note that this expression is consistent with the Einstein summation convention.

An example of an M on euclidean space could be a rotation matrix, or the identity.

When M satisfies 2.1, it is said to be orthogonal. It $\det(M) = 1$ then we say that M is *special*.

2.11 Trace of M

What is the trace of M?

$$\mathrm{Tr}(M) = M^{\alpha}{}_{\alpha} = M^{0}{}_{0} + \ldots + M^{n}{}_{n}$$

This is just a notationally convention. It is the sum of the diagonal terms of M.

2.12 Tensor Product

A tensor product makes a linear map out of a multi-linear map.

Theorem:
Let E and F be 2 vector spaces (with finite dimensionality.)
$\exists$ a unique (!) set (up to isomorphism) $E \otimes F$ such that if f is a bilinear map $f : E \times F \to \mathbb{R}$ then $\exists$ a linear map $f^* : E \otimes F \to \mathbb{R}$ such that $f = f^* \circ \phi$ with

$$\begin{tikzcd} E \times F \arrow[d, "\phi"'] \arrow[dr, "f"] & \\ E \otimes F \arrow[r, "f^*"'] & \mathbb{R} \end{tikzcd}$$

Then we have,

$$\mathrm{Lin}(E \otimes F, \mathbb{R}) \cong \mathrm{Bin}(E \times F, \mathbb{R})$$
$$\mathrm{Lin}(f^*, \mathbb{R}) \cong \mathrm{Bin}(f, \mathbb{R})$$

where '$\cong$' is used to denote *isomorphic*.

Properties:
Basis for $E \otimes F$ is $e_\alpha \otimes g_\alpha$ where e_α is the basis for E and g_α is the basis for F. For $a \in \mathbb{R}$ and $t, v \in E$, $u, w \in F$,

- $\dim(E \otimes F) = \dim(E)\dim(F)$
- $a(v \otimes w) = (av) \otimes w = v \otimes (aw)$
- $(v + t) \otimes w = v \otimes w + t \otimes w$
- $v \otimes (w + u) = v \otimes w + v \otimes u$
- $a \otimes w = aw$
- $\mathbb{R} \otimes F = F$

Note that $V^* \otimes V^* \cong \mathrm{Bin}(V \times V, \mathbb{R})$. To motivate this, let $f^\alpha \otimes f^\beta$ be the basis for $V^* \otimes V^*$, and then a general element in $V^* \otimes V^*$ is,

$$t = t_{\alpha\beta} f^\alpha \otimes f^\beta$$

Note that $t_{\alpha\beta}$ is just a set of numbers. Then the tensor product is expanded as follows,

$$\begin{aligned} t(v \otimes w) &= t(v^\alpha e_\alpha \otimes w^\beta e_\beta) \\ &= t_{\gamma\delta}(f^\gamma \otimes f^\delta)(v^\alpha e_\alpha \otimes w^\beta e_\beta) \\ &= t_{\gamma\delta} v^\alpha w^\beta (f^\gamma \otimes f^\delta)(e_\alpha \otimes e_\beta) \quad \text{By linearity} \end{aligned}$$

$$
\begin{aligned}
&= t_{\gamma\delta} v^\alpha w^\beta f^\gamma(e_\alpha) f^\delta(e_\beta) \quad \text{By foiling and definition of } f \\
&= t_{\gamma\delta} v^\alpha w^\beta \delta^\gamma \alpha \delta^\delta{}_\beta \\
&= t_{\gamma\delta} v^\gamma w^\beta \delta^\delta{}_\beta \quad \text{By sifting property of } \delta \\
&= t_{\gamma\delta} v^\gamma w^\delta \quad \text{By sifting property of } \delta \text{ again}
\end{aligned}
$$

Since $t(v \otimes w)$ is the tensor product $V^* \otimes V^*$ and $t_{\gamma\delta}$ is the components of the bilinear form, one can see the connection $V^* \otimes V^* \cong \mathrm{Bin}(V \times V, \mathbb{R})$.

Tensors allow one to write bilinear maps as linear maps. What about multi-linear maps?

Tensors:
A tensor of rank (k, l) is a multi-linear map

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

which transforms *well* under the change of basis of V and V^*.

Tensor	Rank
vectors	$(1, 0)$
co-vectors	$(0, 1)$
scalar	$(0, 0)$
metric	$(0, 2)$
inverse metric	$(2, 0)$
matrix	$(1, 1)$

The set of tensors of rank (k, l) is a vector space of dimension n^{k+l} (if V has dimension n). Checking with the examples above motivates this fact.

Using the basis $e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \otimes f^{\beta_1} \otimes \cdots \otimes f^{\beta_k}$

$$T = T^{\alpha_1\alpha_2\cdots\alpha_k}{}_{\beta_1\beta_2\cdots\beta_l} e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \otimes f^{\beta_1} \otimes \cdots \otimes f^{\beta_k}$$

For fixed α_i and β_i this is a real number in $\mathbb{R}$. These are the *components of the tensor*.

By abuse of notation we will call $T^{\alpha_1\alpha_2\cdots\alpha_k}{}_{\beta_1\beta_2\cdots\beta_l}$ the tensor.

We are talking about these transformations as change of basis of V and V^*. Examples:

- rotations (boost)
- change of coordinates from Cartesian to spherical, cylindrical, etc.

We can have a linear change of basis $\tilde{x}^\mu = A^\mu{}_\nu x^\nu$.

Example:

Cartesian	Polar
$e_1 = \vec{i}$	$\tilde{e}_1 = e_r$
$e_2 = \vec{j}$	$\tilde{e}_2 = e_\theta$

Example:

$$\tilde{e}_\alpha = \underbrace{\frac{\partial x^\nu}{\partial \tilde{x}^\alpha}}_{\text{Jacobian}} e_\nu = A^\nu{}_\alpha e_\nu$$

Note: *Up in the denominator means down on the original coordinates (LHS).*
For example,

$$\begin{array}{c|c} x^1 = x & \tilde{x}^1 = r \\ x^2 = y & \tilde{x}^2 = \theta \end{array}$$

$$\tilde{e}_1 = e_r = \frac{\partial x^1}{\partial \tilde{x}^1} e_1 + \frac{\partial x^2}{\partial \tilde{x}^1} e_2 = \cos\theta e_1 + \sin\theta e_1$$

$$\tilde{e}_2 = e_\theta = \frac{\partial x^1}{\partial \tilde{x}^2} e_1 + \frac{\partial x^2}{\partial \tilde{x}^2} e_2 = -r\sin\theta e_1 + r\cos\theta e_1$$

Vectors in multiple basis:

$$v = v^\nu e_\nu = \tilde{v}^\nu \tilde{e}_\nu$$

With conversion of basis given by,

$$\tilde{e}_\alpha = A^\nu{}_\alpha e_\nu$$

Thus substituting in,

$$v^\nu e_\nu = \tilde{v}^\alpha A^\nu{}_\alpha e_\nu \quad \text{Drop } e_\nu$$

$$v^\nu = \tilde{v}^\alpha A^\nu{}_\alpha$$

But with A as a Jacobian,

$$v^\nu = \frac{\partial x^\nu}{\partial \tilde{x}^\alpha} \tilde{v}^\alpha$$

$$\tilde{v}^\alpha = \frac{\partial \tilde{x}^\alpha}{\partial x^\nu} v^\nu$$

But what about the dual space?
By definition,

$$\tilde{f}^\beta(\tilde{e}_\nu) = \delta^\beta_\mu = \tilde{f}^\beta(A^\alpha{}_\nu e_\alpha) = A^\alpha{}_\nu \tilde{f}^\beta(e_\alpha)$$

Let $\tilde{f}^\beta(e_\alpha)$ be expressed as $\tilde{f}^\beta = B^\beta{}_\gamma f^\gamma$

$$\begin{aligned} \tilde{f}^\beta(\tilde{e}_\nu) &= A^\alpha{}_\nu B^\beta{}_\gamma f^\gamma(e_\alpha) \\ &= A^\alpha{}_\nu B^\beta{}_\gamma \delta^\gamma{}_\alpha \\ &= B^\beta{}_\gamma A^\gamma{}_\nu \\ &= \delta^\beta{}_\nu \end{aligned}$$

Thus B is the inverse of A.

What does transforming *well* mean? A tensor is transforming well if its components transform as

$$T^{\nu_1\nu_2\cdots\nu_k}{}_{\alpha_1\alpha_2\cdots\alpha_l} \to \frac{\partial \tilde{x}^{\nu_1}}{\partial x^{\beta_1}} \cdots \frac{\partial \tilde{x}^{\nu_k}}{\partial x^{\beta_k}} \frac{\partial x^{\gamma_1}}{\partial \tilde{x}^{\alpha_1}} \cdots \frac{\partial x^{\gamma_k}}{\partial \tilde{x}^{\alpha_k}} T^{\beta_1\beta_2\cdots\beta_k}{}_{\gamma_1\gamma_2\cdots\gamma_l} = \tilde{T}^{\nu_1\nu_2\cdots\nu_k}{}_{\alpha_1\alpha_2\cdots\alpha_l}$$

If you find something like $T^\alpha{}_\beta$, is it a tensor? **No! You must check if it transforms well.**

$$\frac{\partial}{\partial x^\nu} v^\alpha \quad \text{This is not a tensor.}$$

The derivative here prevents it from being well-formed. In the future we will define a derivative that allows a tensor to transform well.

2.13 Operations on Tensors

- Add (with matching rank): $T^{\alpha_1\alpha_2}{}_{\beta_1\beta_2} + C^{\alpha_1\alpha_2}{}_{\beta_1\beta_2}$.
- Contraction (partial trace): $\mathcal{T}(k,k) \to \mathcal{T}(k-1,k-1)$.
  - $T^{\alpha_1\cdots\alpha_i\cdots\alpha_k}{}_{\beta_1\cdots\beta_j\cdots\beta_l} \to T^{\alpha_1\cdots\alpha_i\cdots\alpha_k}{}_{\beta_1\cdots\alpha_j\cdots\beta_l}$
- "Outer" Product (Gluing together tensors)
  - $\mathcal{T}(k,l) \times \mathcal{T}(k',l') \to \mathcal{T}(k+k',l+l')$
  - $(T_1,T_2) \to T_1T_2$
  - $T_1T_2 \to T_1^{\nu_1\cdots\nu_k}{}_{\alpha_1\cdots\alpha_l} T_2^{\beta_1\cdots\beta_k}{}_{\gamma_1\cdots\gamma_l}$
  - **Example:** $(v^\alpha, w_\beta) \to v^\alpha \otimes w_\beta = v^\alpha w_\beta$. (In QM this is $|\phi\rangle\langle\varphi|$)

The metric $g_{\alpha\beta}$ can change the rank of a tensor. Recall a metric is rank $(0,2)$ is symmetric and is non-degenerate.

Example:
Changing from rank $(1,0)$ to rank $(0,1)$:

$$v^\alpha \to v_a = g_{\alpha\beta}v^\beta$$

Changing from rank $(2,2)$ to rank $(4,0)$:

$$C^{\alpha\beta}{}_{\gamma\delta} \to C_{\alpha\beta\gamma\delta} = g_{\alpha\rho}g_{\beta\eta}C^{\rho\eta}{}_{\gamma\delta}$$

Changing from rank $(2,2)$ to a different rank $(2,2)$:

$$C^{\alpha\beta}{}_{\gamma\delta} \to C^{\alpha}{}_{\beta}{}^{\gamma}{}_{\delta} = g_{\beta\rho}g^{\gamma\eta}C^{\alpha\rho}{}_{\eta\delta}$$

2.14 Facts About Tensors

Order Matters:
The order of indices that label a tensor is **very** important. It indicates the product space you are mapping *from* to $\mathbb{R}$.

$$\begin{aligned} C^{\alpha}{}_{\beta}: &\quad V^* \times V \to \mathbb{R} \\ C_{\alpha}{}^{\beta}: &\quad V \times V^* \to \mathbb{R} \\ C^{\beta}_{\alpha}: &\quad \text{Nothing. Don't do this.} \end{aligned}$$

Equality between tensors:
As tensors, indices must match:

Position of indices is matching: $C^{\alpha}{}_{\gamma}{}^{\delta} = T^{\alpha}{}_{\gamma}{}^{\delta}$

Position of indices is **not** matching: $C^{\alpha}{}_{\gamma}{}^{\delta} \neq T^{\alpha}{}_{\gamma\delta}$

But for fixed α,γ,δ, one can abuse the notation a bit:

$$C^{\alpha}{}_{\gamma}{}^{\delta} = T^{\alpha}{}_{\gamma\delta} \quad \text{Try to avoid this.}$$

2.15 Outer Product and Contraction

Example:
Outer Product: $M^{\alpha}{}_{\beta} M^{\gamma}{}_{\delta} = C^{\alpha}{}_{\beta}{}^{\gamma}{}_{\delta}$
Contraction: $M^{\alpha}{}_{\beta} M^{\beta}{}_{\delta} = C^{\alpha}{}_{\beta}{}^{\beta}{}_{\delta} = C^{\alpha}{}_{\delta}$

Example:
Outer product and contraction: $C^{\alpha\beta}{}_{\gamma\delta} T^{\gamma\delta}{}_{\rho} = A^{\alpha\beta}{}_{\rho}$
This doesn't make sense: $C^{\alpha\beta}{}_{\gamma\gamma} T^{\gamma\delta}{}_{\rho} = ??$

Note, when there is a "+" sign we can be "loose" with the indices. Here the dual indices **do not** indicate a summation. This acts as an abuse of notation, but is sometimes difficult to avoid.

$$C^{\alpha\gamma} T_{\gamma}{}^{\delta} + F_{\gamma}{}^{\delta} A^{\alpha\gamma}$$

2.16 Interpretation of Tensors

By looking at the indices, how can we interpret the physical meaning of the tensor object?

Tensor	Interpretation
v^{ν}	vector
v_{ν}	covector
$M^{\alpha}{}_{\beta}$	matrix (α rows, β columns)
$M^{\alpha}{}_{\alpha}$	contracted matrix (trace)
$M^{\alpha\gamma}{}_{\delta}$	matrix whose elements are vectors themselves ($\cdot^{\gamma}{}_{\delta}$ is the matrix)
$M^{\alpha\gamma}{}_{\delta}$	vector with matrix components (M^{α} is the vector)
$R^{\alpha\beta}{}_{\gamma\delta}$	matrix of matricies *

*For example, if $\dim V = 4$, $R^{\alpha\beta}{}_{\gamma\delta}$ has $4^4 = 256$ components. Note however, there can be many symmetries that reduce the number of unique components.

2.17 Symmetry of Tensor

We can always build a symmetric and antisymmetric part of a tensor $T^{\alpha\beta}$. Let's look at the case of 2 indices α, β:

Symmetric Part:

$$T_{(\alpha\beta)} = \frac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha})$$
$$T_{(\alpha\beta)} = T_{(\beta\alpha)}$$

Antisymmetric Part:

$$T_{[\alpha\beta]} = \frac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha})$$
$$T_{[\alpha\beta]} = -T_{[\beta\alpha]}$$

Note that for all tensors $T^{\alpha\beta} = T^{(\alpha\beta)} + T^{[\alpha\beta]}$. This acts as the decomposition into odd and even symmetries of the tensor.

For more indices:

$$T^{(\alpha\beta)}{}_{[\gamma\delta]} = \frac{1}{4}\left(T^{\alpha\beta}{}_{\gamma\delta} + T^{\beta\alpha}{}_{\gamma\delta} - T^{\alpha\beta}{}_{\delta\gamma} - T^{\beta\alpha}{}_{\delta\gamma}\right)$$

What does $T^{(\alpha\beta\gamma)}$ mean? For that we will need a permutation group.

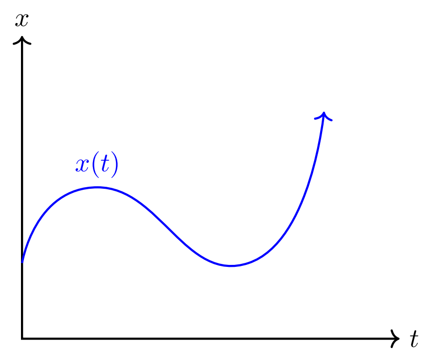
3 Physics Review

Moving away from tensors for a moment...

3.1 Newtonian Physics

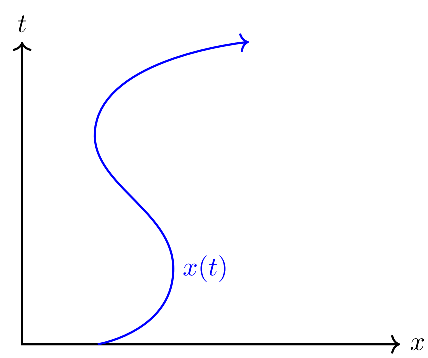
According to Galileo and Newton, we got the interpretation that both space and time is flat $\left(\mathbb{R}^3\right)$ and is absolute. More specifically, all clocks will have the same time if they are started/synced at some shared moment. It is built on cartesian coordinate system: $(\vec{x}, t)$. With this we say that an object is at position $\vec{x}$ at time t. In this context, coordinates are *outcomes of measurements*. In General Relativity, the notion of coordinates can be quite different.

Consider a particle (2d spacetime):



Typically, x is drawn as the ordinate (y-axis) and t as the abscissa (x-axis).

Spacetime diagram:
In a spacetime diagram, t is drawn as the ordinate.



If we begin to use light to probe the position of objects, we are going to run into some surprising results. We will have to abandon Newtonian Physics and switch to the domain of Special Relativity.

3.1.1 Newton's Dynamical Law & Inertial Observers

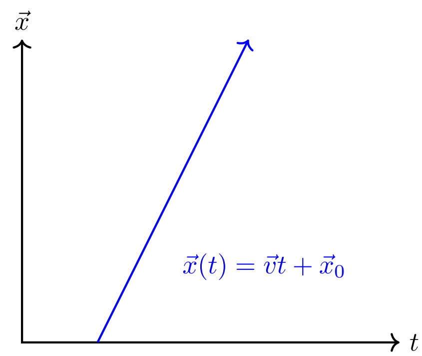
$$\vec{F} = m\vec{a}$$

Where $\vec{F}$ is the total force applied to the system, $\vec{a} = \ddot{\vec{x}} = \frac{\mathrm{d}^2\vec{x}}{\mathrm{d}t^2}$ and m is the inertial mass. For $\vec{x}$ is it convienent to use the Cartesian coordinate system.

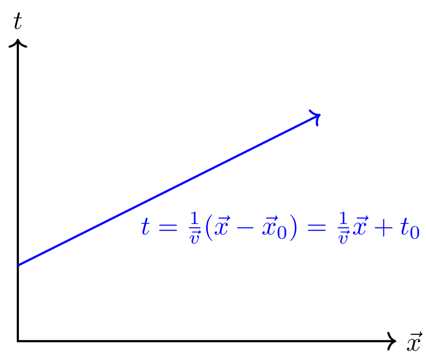
If $\vec{F} = \vec{0}$ then the dynamics becomes $\ddot{\vec{x}} = 0$ which yields solution,

$$\vec{x}(t) = \vec{v}t + \vec{x}_0$$

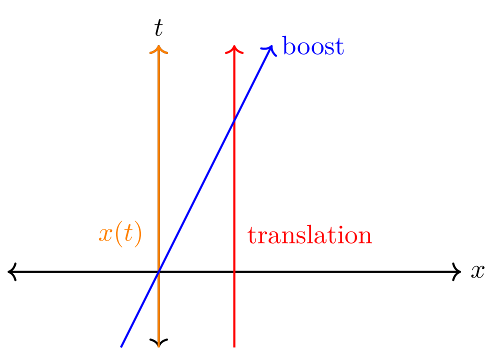
Where $\vec{x}_0$ is the initial condition and $\vec{v}$ is the velocity in the observer's frame. This solution describes a straight line.



However, consider this solution in a spacetime diagram,



Definition:
The class of frames (observers) for which the the dynamics of a system is $\ddot{\vec{x}} = \vec{0}$ are called an inertial observers.



Note that rotations are not visible in this diagram as there is only one 1 space dimension.
Transformations that relate inertial observers:

- translation: $\vec{x}(t) \to \vec{x}\,'(t) = \vec{x}(t) + \vec{a}$
- rotation: $\vec{x}(t) \to \vec{x}\,'(t) = R \cdot \vec{x}$
- Galilean boost: $\vec{x}(t) \to \vec{x}\,'(t) = -\vec{v}t + \vec{x}(t)$

For each of these transformations $\ddot{\vec{x}}' = \vec{0}$. We will now prove the set of all these transformation of $\vec{x}(t) \to \vec{x}\,'(t)$ form a group.

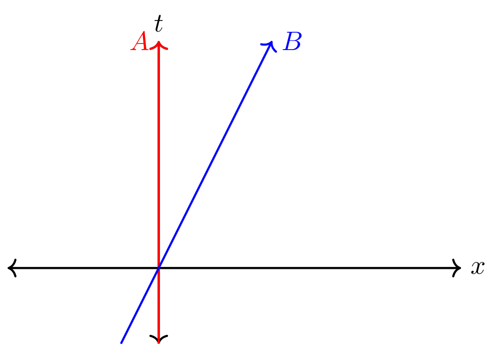
Groups:
A Group is a set G equipped with an associated product $(\cdot)$, a unit element and an inverse.

- $g_1 \cdot g_2 = g \in G, g_i \in G$
- $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$
- $g \cdot 1 = 1 \cdot g = g$
- $g \cdot g^{-1} = g^{-1} \cdot g = 1$
- In general, $g_1 \cdot g_2 \neq g_2 \cdot g_1$.
- An abelian group is one where $g_1 \cdot g_2 = g_2 \cdot g_1$.

Upon careful examinations of translations and rotation of space $\mathbb{R}^n$, both translations and rotations form a group. What about Galilean boosts?

Consider person A (Alice) standing on the ground and person B (Bob) in a rocket traveling with velocity $\vec{v}_1$ with respect to B. Give person B a ball in the rocket and let him/her kick it with velocity $\vec{v}_2$ with respect to B. What is the velocity of the ball with respect to person A? Switching between the perspectives of the system is equivalent to performing a Galilean Boost.

Matrix Representation of a Group:



The boost $B^{\alpha}{}_{\gamma}$ is given by,

$$B^{\alpha}{}_{\gamma} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -v_1 & 1 & 0 & 0 \\ -v_2 & 0 & 1 & 0 \\ -v_3 & 0 & 0 & 1 \end{bmatrix}$$

Person A (sitting on the ground) is given by,

$$A \sim \begin{bmatrix} t \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Their product is given by,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -v_1 & 1 & 0 & 0 \\ -v_2 & 0 & 1 & 0 \\ -v_3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} t \\ -v_1 t \\ -v_2 t \\ -v_3 t \end{bmatrix}$$

This is the trajectory of Alice with respect to Bob parametrized with time t. What about Bob's perspective under this linear map?

$$\begin{bmatrix} t \\ v_1 t \\ v_2 t \\ v_3 t \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -v_1 & 1 & 0 & 0 \\ -v_2 & 0 & 1 & 0 \\ -v_3 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} t \\ -v_1 t + v_1 t \\ -v_2 t + v_2 t \\ -v_3 t + v_3 t \end{bmatrix} = \begin{bmatrix} t \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus Galilean boosts form an abelian group. When we move to the regime of Special Relativity, we will see that boosts no longer form an abelian group.

3.2 The Relativity Principle

Two inertial observers moving with constant velocity cannot be distinguished by any physical experiment.

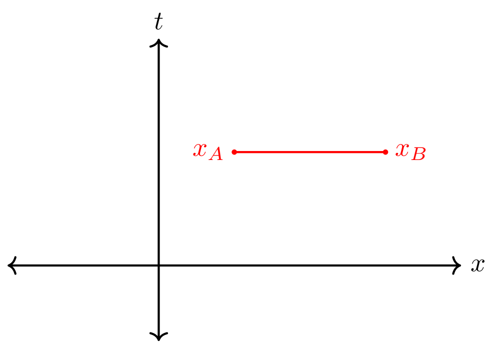
Or alternatively,

Inertial frames are equivalent in terms of the description of physical phenomena.

This is most easily observed when sitting on a train next to another train. When your train is moving, it is unclear whether or not your train is moving or the other one is.

Inertial frames are systems with $\ddot{\vec{x}} = \vec{0}$ equipped with rods and clocks for measurements.

What happens to the notion of spatial length when you change inertial frames? Nothing should change due to the Relativity Principle, the lengths should remain the same.



In one frame,

$$|\vec{x}_B - \vec{x}_A|^2 = \ell^2$$

Perform a Galilean boost,

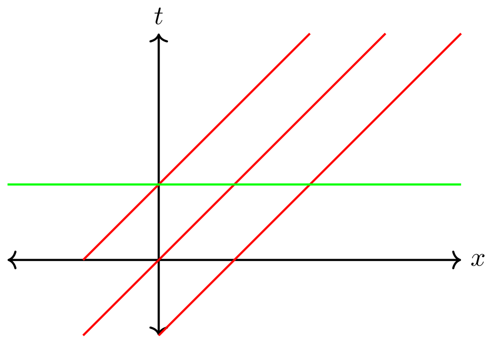
$$\vec{x}'_{A,B} = -\vec{v}t + \vec{x}_{A,B}$$

$$|\vec{x}'_B - \vec{x}'_A|^2 = \ell'^2$$

But it must be that,

$$\ell' = \ell$$

What about time? Galilean transformation leave time invariant because time is absolute in this Newtonian regime. The notion of simultaneity is the same for any inertial observer.



Red lines are stationary observers, and intersection with red lines indicate simultaneous events.

Math Perspective:
Are Galilean transformations the most general transformations between inertial observers? Use axioms?

Physics Perspective:

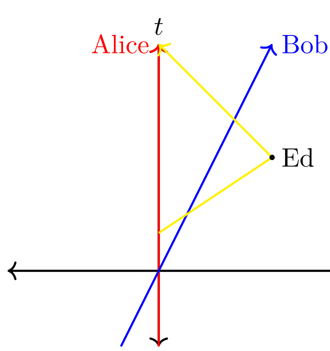
- Maxwell's equations do not transform well under Galilean transforms. Lorentz found the Lorentz transformations that allow Maxwell's equations to transform well.
- Michelson-Morley experiment: reveals that speed of light is invariant under the change of frame.
  - $v_1 + v_2 = v_3$ This is **not** the case if $v_2 = c$
  - $v_1 + c = c$ What why???
  - Under the assumption that light is a wave in the ether. Results suggest that the ether is not measurable.

3.3 Lorentz Transformations

Let's use light to measure objects in two frames; specifically let's determine the position using light.

Assumptions:

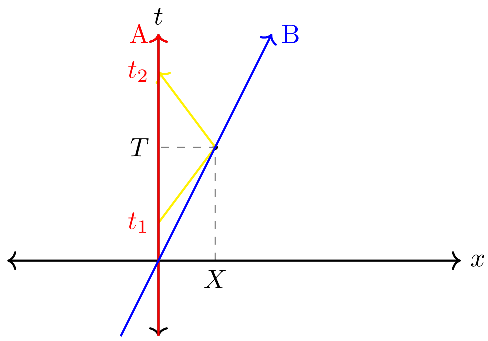
- The speed of light is the same in any frame.
- Relativity Principle
- Bob will move at velocity $v < c$
  - If an observer is moving $v > c$, their position can't be measured using light
- 2d for simplicity
- Set $c = 1$, $x = ct + x_0 = t + x_0$



Let's measure the coordinates of Ed in Alice's frame and Bob's frame and the map relating the times using light to measure positions.
1) Using light to measure position,

$$X = \frac{1}{2c}(t_2 - t_1)$$

$$T = t_1 + \frac{1}{2}(t_2 - t_1) = \frac{1}{2}(t_2 + t_1)$$



2) Determine how the difference of times of reception and emission are related?

$$\Delta t_A = ON \quad \Delta t_B = QP$$

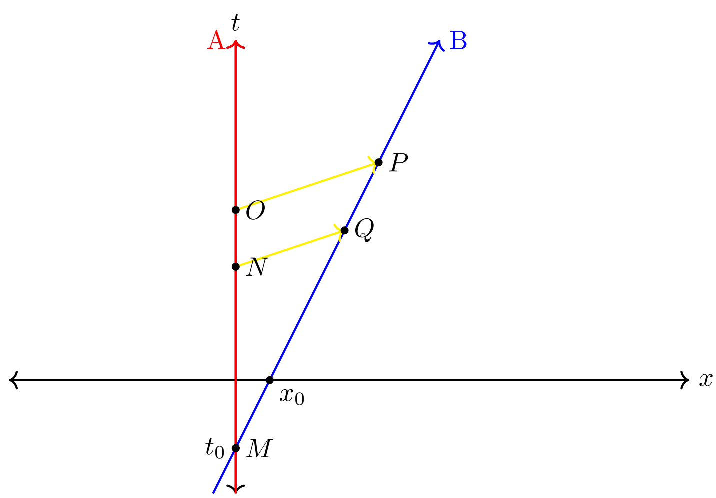
$$\frac{MO}{MN} = \frac{MP}{MQ}$$

$$\frac{MN + ON}{MN} = \frac{MQ + QP}{MQ}$$

$$\frac{ON}{MN} = \frac{QP}{MQ}$$

$$\frac{\Delta t_A}{MN} = \frac{\Delta t_B}{MQ}$$

$$\Delta t_A \propto \Delta t_B$$



$$\Delta t_A = f^{-1}(v, c, x_0, t_0)\Delta t_B$$

By translational invariance, f cannot depend on x_0 or t_0. Therefore,

$$\Delta t_A = f^{-1}(v, c)\Delta t_B$$

For convenience, we will find f defined as,

$$\Delta t_B = f(v, c)\Delta t_A$$

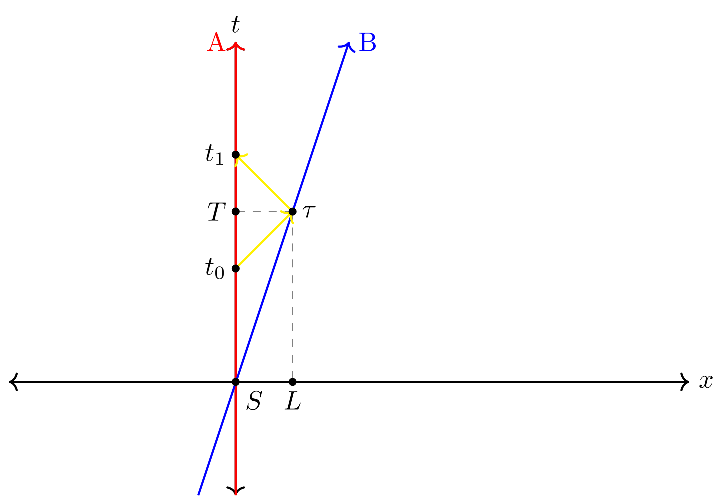
By similar analysis of a pair of light rays emanating from bob,

$$\Delta \tilde{t}_A = \tilde{f}(v, \tilde{c})\Delta \tilde{t}_B = \tilde{f}(v, c)\Delta \tilde{t}_B$$

This uses the assumption $\tilde{c} = c$. Furthermore by the relativity principle, we must have $\tilde{f} = f$. Otherwise, the two inertial frames (A, B) would be distinguishable through $f, \tilde{f}$. In conclusion:

$$\Delta t_{\text{received}} = f(v, c)\Delta t_{\text{emitted}}$$

So what is the form of f? Let us assume that there is a synchronization between the two frames so that calculations become easier.



Let's define:

$$\begin{aligned}
\Delta t_A &= S \to t_0 = t_0 \\
\Delta \tilde{t}_A &= S \to t_1 = t_1 \\
\Delta t_B &= S \to \tau = \tau = \Delta \tilde{t}_B
\end{aligned}$$

Therefore $\tau = f(v, c)t_0$ with,

$$t_1 = \Delta \tilde{t}_A = f(v, c)\Delta \tilde{t}_B = f^2(v, c)t_0$$

Thus we have from radar measurements,

$$L = \frac{1}{2}c(t_1 - t_0) = \frac{1}{2}c\big(f^2(v, c) - 1\big)t_0$$

$$T = \frac{1}{2}(t_1 + t_0) = \frac{1}{2}c\big(f^2(v, c) + 1\big)t_0$$

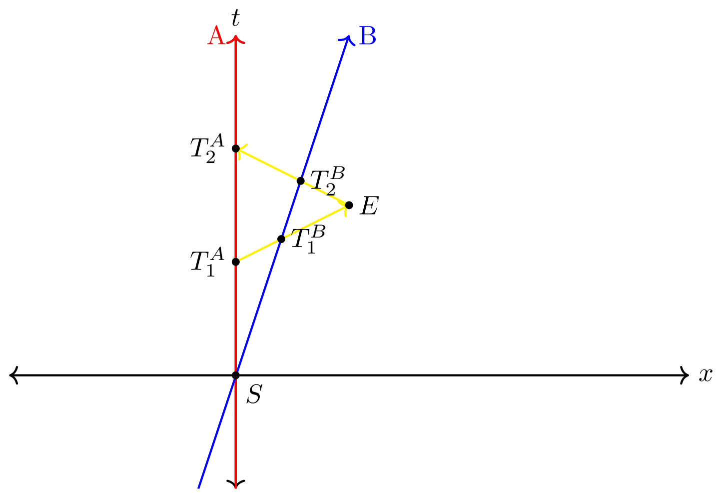
The ratio is given by,

$$\frac{L}{T} = v = c\frac{f^2(v,c)-1}{f^2(v,c)+1}$$

Inverting this expression (using $-c < v < c$) yields,

$$f(v,c) = \left(\frac{1+\frac{v}{c}}{1-\frac{v}{c}}\right)^{1/2}$$

4) Ed's coordinate from A's and B's perspective.



Let E be identified in two ways,

Alice's Perspective	(x_A^E, t_A^E)
Bob's Perspective	(x_B^E, t_B^E)

Therefore,

$$x_A^E = \frac{1}{2}c\big(T_2^A - T_1^A\big)$$
$$t_A^E = \frac{1}{2}\big(T_2^A + T_1^A\big)$$
$$x_B^E = \frac{1}{2}c\big(T_2^B - T_1^B\big)$$
$$t_B^E = \frac{1}{2}\big(T_2^B + T_1^B\big)$$

We know that $T_1^B = fT_1^A$ and $T_2^A = fT_2^B$ hence we can get the relation between (x_A^E, t_A^E) and (x_B^E, t_B^E).

$$x_B^E = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\big(x_A^E - vt_A^E\big)$$
$$t_B^E = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(t_A^E - \frac{v}{c^2}x_A^E\right)$$

For convenience we can relabel,

$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

In matrix form,

$$\begin{bmatrix} T' \\ X' \end{bmatrix} = \underbrace{\gamma\begin{bmatrix} 1 & \frac{-v}{c^2} \\ -v & 1 \end{bmatrix}}_{\text{Lorentz Boost}}\begin{bmatrix} T \\ X \end{bmatrix}$$

Notice in the limit that $v << c$, the Lorentz boost becomes equivalent to the Galilean boost discussed earlier.

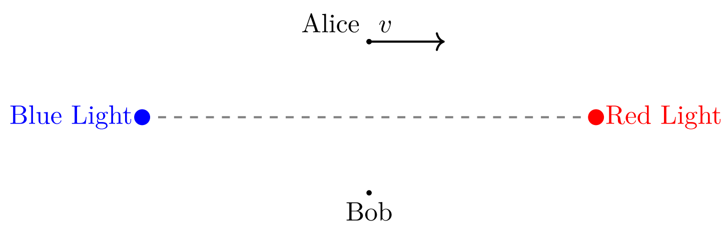
3.3.1 Consequences of Lorentz Transformations

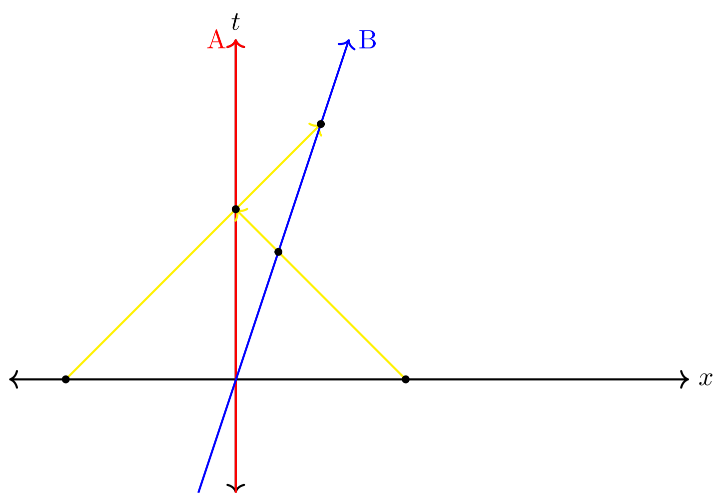
1. From Alice's perspective, Bob's time axis is,
   - $x^B = 0 = \gamma(x_A - vt_A) \implies x = vt$
2. From Alice's perspective, Bob's spatial slice is,
   - $T^B = 0 = \gamma\big(t_A - \frac{v}{c^2}\big)x_A \implies t = \frac{v}{c^2}x$
   - The slope of this line in Alice's perspective is the **inverse** (with $c \to 1$) of the slope of Bob's time axis

Again, notice that Galilean simultaneity is recovered in the limit that $v \to v \ll c$. Our new theory is still consistent with our old theories.

3.4 Length & Time

3.4.1 Einstein's Train & Simultaneity





Therefore, simultaneity is **relative**.

3.4.2 Length

What about spatial length? Alice has a ruler of length $x_2 - x_1 = \ell$. In Bob's frame,

$$x_i = \left(\underbrace{x_i' + vt'}_{\text{Bob's coords}}\right)\gamma \quad i = 1, 2$$

Thus,

$$x_2' - x_1' = (x_2 - x_1)\frac{1}{\gamma} \implies \ell' = \frac{\ell}{\gamma} \quad \text{Length contraction.}$$

If Bob has a ruler of length $\ell' = x_2' - x_1'$,

$$x_i' = \left(\underbrace{x_i + vt}_{\text{Alice's coords}}\right)\gamma \quad i = 1, 2$$

$$x_2 - x_1 = \ell = \frac{1}{\gamma}\ell' \quad \text{Length contraction.}$$

3.4.3 Time

In a similar manner, we have time dilation,

$$t_2' - t_1' = \gamma(t_2 - t_1) \quad \text{With } \gamma > 1.$$

3.4.4 Invariant Length & Minkowski Metric

Can we construct/define a notion of length that is invariant under these transformations? Namely,

$$\begin{bmatrix} \gamma & -\gamma\frac{v}{c^2} \\ -\gamma v & \gamma \end{bmatrix} = \Lambda^{\alpha}{}_{\beta}$$

Let's introduce a metric, $g_{\alpha\beta}$ that is of course symmetric and non non-degenerate. Therefore,

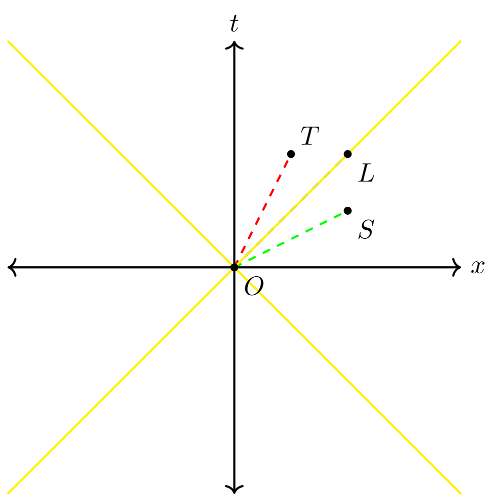
$$\Lambda^{\alpha}{}_{\beta} g_{\alpha\gamma} \Lambda^{\gamma}{}_{\delta} = g_{\beta\delta}$$

Reminder: $v \cdot w = v' \cdot w'$ with $v'^{\alpha} = \Lambda^{\alpha}{}_{\gamma} v^{\gamma}$. The solution here for 1d motion:

$$g_{\alpha\beta} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

These are called *Minkowski Metrics*. The difference between these two is a matter of convention. We will select the convention,

$$g_{\alpha\beta} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$



Let us define,

$$\left|\vec{OT}\right|^2 < 0 \quad \text{“time like”}$$

$$\left|\vec{OL}\right|^2 = 0 \quad \text{“light like”}$$

$$\left|\vec{OS}\right|^2 > 0 \quad \text{“space like”}$$

We say that a vector v^α is time-like if $v^\alpha \eta_{\alpha\beta} v^\beta \equiv |v|^2 < 0$, light-like if $v^\alpha \eta_{\alpha\beta} v^\beta \equiv |v|^2 = 0$ and space-like if $v^\alpha \eta_{\alpha\beta} v^\beta \equiv |v|^2 > 0$.

What are inertial observers in the relativistic regime? They are *still* given by a *straight lines*. What are the set of transformations relating inertial observers (4d) ?

- rotations: R_x, R_y, R_z

$$\to \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- boost: B_x, B_y, B_z

$$B_z \to \begin{bmatrix} \gamma & -\gamma\frac{v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore,

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

Let us relabel $\gamma^2 = \cosh^2\eta$ and $\gamma^2 \frac{v^2}{c^2} = \sinh^2\eta$,

$$B_z \to \begin{bmatrix} \cosh\eta & -\frac{\sinh\eta}{c^2} & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- translation: $\vec{x} \to \vec{x} + \vec{a}$; $t \to t + b$
  - $x^\nu \to x^\nu + a^\nu$

All of these together form the Poincaré group.

3.4.5 Poincaré Group

Galilean boosts:

$$\vec{v}_1 + \vec{v}_2 = \vec{v}_3$$

Lorentzian boosts:

$$\overrightarrow{v_1 \oplus v_2} = \frac{1}{1 + \frac{\vec{v}_1 \cdot \vec{v}_2}{c^2}} \left(\vec{v}_1 + \frac{1}{\gamma_1}\vec{v}_2 + \frac{1}{c}\frac{\gamma_1}{1-\gamma_1}(\vec{v}_1 \cdot \vec{v}_2)\vec{v}_2 \right)$$

Note these aren't associative,

$$\overrightarrow{\overrightarrow{v_1 \oplus v_2} \oplus v_3} \neq \overrightarrow{v_1 \oplus \overrightarrow{v_2 \oplus v_3}}$$

However it $\vec{v}_1$ is parallel to $\vec{v}_2$ then,

$$\overrightarrow{v_1 \oplus v_2} = \frac{\vec{v}_1 + \vec{v}_2}{1 + \frac{\vec{v}_1 \cdot \vec{v}_2}{c^2}}$$

Which forces $\left|\overrightarrow{v_1 \oplus v_2}\right|^2 = c^2$ whenever either $\vec{v}_1$ or $\vec{v}_2$ has $|\vec{v}_i|^2 = c^2$.

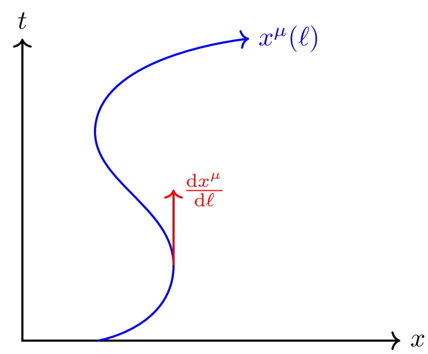
4 Formulation of Gravity

4.1 Flat Spacetime

Spacetime is $\mathbb{R}^4$ and we will consider the Minkowski metric $\eta_{\alpha\beta}$. In Cartesian coordinates,

$$x^\mu = \begin{bmatrix} ct & x & y & z \end{bmatrix}$$

Consider a curve in spacetime $x^\nu(\ell)$ where ℓ is a curvilinear parameter that parametrizes the curve. The tangent vector is given by $\frac{\mathrm{d}x^\mu}{\mathrm{d}\ell}$.



If the tangent $\frac{\mathrm{d}x^\mu}{\mathrm{d}\ell}$ is always time-like, then the curve is called time-like. Analogously for space-like and light-like curves. We can then define an arclength of the curve.

$$\tau \neq \int \sqrt{\frac{\mathrm{d}x^\mu}{\mathrm{d}\ell} n_{\mu\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\ell}} \mathrm{d}\ell$$

This expression does not work because the metric $\eta_{\mu\nu}$ is not positive definite. Therefore the length could be complex. To combat this, introduce an absolute magnitude,

$$\tau = \int \sqrt{\left|\frac{\mathrm{d}x^\mu}{\mathrm{d}\ell} n_{\mu\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\ell}\right|} \mathrm{d}\ell$$

Which suggests,

$$\frac{\mathrm{d}\tau}{\mathrm{d}\ell} = \sqrt{\left|\frac{\mathrm{d}x^\mu}{\mathrm{d}\ell} n_{\mu\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\ell}\right|}$$

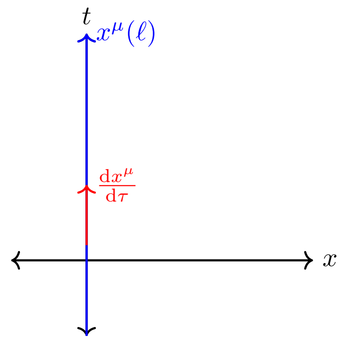
Which implies

$$\mathrm{d}\tau^2 = |\mathrm{d}x^\mu n_{\mu\nu} \mathrm{d}x^\nu|$$

Now consider using τ as the curvilinear parameter. So that the tangent is defined as,

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$$

What does this encapsulate? Consider an observer at rest:



Here, the notion of arclength coincides with time.

$$d\tau^2 = dt^2$$

Considering a time-like curve, the proper length or arclength is interpreted as the time as measured by a clock carried by the observer having the curve as his/her spacetime trajectory (world-line).

Parametrization of the world line of an observer at rest ($\vec{x} = \vec{0}$).

$$x^0(\tau) = \tau \quad \text{and} \quad \vec{x} = \vec{0}$$

Which gives,

$$\frac{dx^\mu}{d\tau} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \delta^\mu{}_0 \neq \delta^0{}_\mu \quad \text{Be careful to matching indices.}$$

Labeling $\delta^\mu{}_0 = V^\alpha$,

$$V^\alpha \eta_{\alpha\beta} V^\beta = \delta^\alpha{}_0 \eta_{\alpha\beta} \delta^\beta{}_0 = \eta_{00} = -1 < 0$$

Therefore, V^α is time-like. However in general,

$$\begin{aligned} V^\mu \eta_{\mu\beta} V^\beta &= \frac{dx^\mu}{d\tau} \eta_{\mu\alpha} \frac{dx^\alpha}{d\tau} \\ &= \frac{dx^\mu \eta_{\mu\alpha} dx^\alpha}{d\tau^2} \\ &= \frac{ds^2}{d\tau^2} \quad ds^2 \text{ is the line element.} \\ &= -1 \quad \text{Since } d\tau^2 = \left|ds^2\right| = -ds^2 \text{ (this vector is time-like).} \end{aligned}$$

We can can also write (with $c = 1$),

$$\begin{aligned} d\tau^2 &= dt^2 - d\vec{x}^2 \\ &= dt^2 \left(1 - \frac{d\vec{x}^2}{dt^2}\right) \\ &= dt^2 \left(1 - \vec{v}^2\right) \end{aligned}$$

Which when rearranged yields (noting the negative is ignored because we are considering time moving forward),

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1 - \vec{v}^2}} = \gamma$$

Noting that,

$$v^\mu = \begin{bmatrix} \frac{\mathrm{d}t}{\mathrm{d}\tau}c \\ \frac{\mathrm{d}\vec{x}}{\mathrm{d}\tau} \end{bmatrix} = \begin{bmatrix} \gamma c \\ \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} \gamma c \\ \gamma \vec{v} \end{bmatrix}$$

Noting that $\vec{v}$ is a velocity in 3d space while v^μ is a relativistic velocity which is a 4d object. We have shown that,

$$v^\mu v_\mu = -1$$

In the Galilean regime, $\vec{v} \in \mathbb{R}^3$ which allows one to use the vector space structure of $\mathbb{R}^3$ to add velocities,

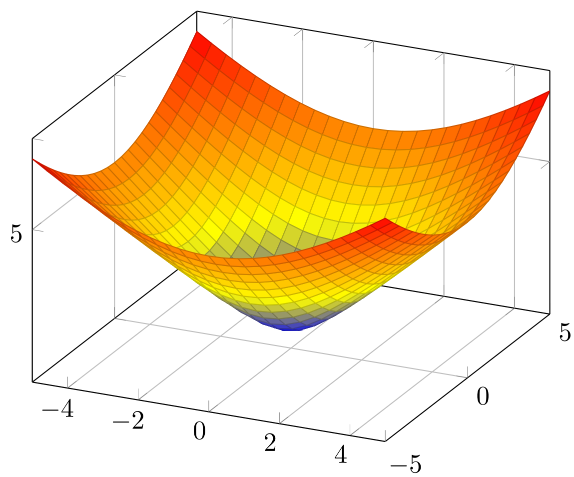
$$\vec{v}_1 + \vec{v}_2 = \vec{v}_{12}$$

But for relativistic velocities $\vec{v} \in \mathbb{H}^3$ hyperboloid.

$$\left\{ v^\mu \in \mathbb{R}^4 \mid v^\mu v_\mu = -1 \right\}$$

This is somewhat surprising, bur there exists a constraint that reduces the dimensionality.

$$v^\mu v_\mu = -1 = -\left(v^t\right)^2 + \left(v^x\right)^2 + \left(v^y\right)^2 + \left(v^z\right)^2$$



This new space $\mathbb{H}^3$ allows for the addition of velocities,

$$\overrightarrow{v_1 \oplus v_2} = \vec{v}_{12}$$

Particle Considerations:

- The world-line of a material particle is a time-like curve
- The world-line of photons (massless) is a light-like curve
- The world-line of a tachyon ($|v| > c$) is space-like

Relativistic acceleration is given by,

$$a^\mu = \frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2}$$

For inertial observer, $a^\mu = 0$.
Relativistic momentum is given for a material particle with mass m,

$$p^\mu = m v^\mu \implies p^\mu p_\mu = -m^2$$

Which has the property,

$$p^\mu = \begin{bmatrix} E = m\gamma \\ \vec{p} = m\gamma\vec{v} \end{bmatrix}$$

Examine the space term,

$$E = m\gamma = m\frac{1}{\sqrt{1-v^2}} = m + \underbrace{\frac{1}{2}mv^2}_{\text{kinetic energy}} + \cdots \quad \text{Taylor Series}$$

And also rest,

$$E^2 = m^2c^4$$

4.2 Relativistic Dynamics

What about the dynamics in the relativistic regime? Recall in the non-relativistic regime, Newton's law is given by,

$$m_i\vec{a} = \vec{F}_{\text{tot}}$$

Where m_i is the inertial mass. How can we modify this equation to the relativistic regime. m_i has no need to change, $\vec{a}$ becomes the 4d spacetime vector $a^\mu = \frac{\mathrm{d}^2x^\nu}{\mathrm{d}\tau^2}$ and force becomes,

$$m_i a^\mu = F^\mu$$

4.2.1 Lorentz Force

Consider the Lorentz force,

$$\vec{F} = q\left(\vec{E} + \vec{v}\times\vec{B}\right)$$

Where q is the charge of the particle, $\vec{E}$ is the electric field, $\vec{v}$ is the velocity of the particle and $\vec{B}$ is the magnetic field of the particle. In the relativistic regime,

$$F^\mu = qF^{\alpha\beta}V_\beta$$

Where $F^{\alpha\beta}$ is known as the Maxwell tensor that is anti-symmetric,

$$F^{\alpha\beta} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{bmatrix}$$

And V_β is the relativistic velocity with,

$$V_\beta = \eta_{\alpha\beta}V^\alpha$$

4.2.2 Gravity

What about the relativistic behavior of gravity?

$$\vec{F} = -m_g\vec{\nabla}\phi = -\frac{Gm_gM\hat{r}}{r^2}$$

Where m_g is gravitational mass, ϕ is the gravitational potential, G is Newton's constant, and M is the mass of the system that generates the gravitational force. For a mass density ρ,

$$M = \iiint_V \rho\,\mathrm{d}V$$

Given a source with mass density ρ we have ϕ given by the Poisson Equation,

$$\vec{\nabla}\vec{\nabla}\phi = \Delta\phi = 4\pi G\rho$$

Where Δ is given by,

$$\Delta = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$$

Gravity tells me how the matter source propagate,

$$\vec{F}_g = m_i\vec{a} \quad \text{gravity} \implies \text{matter} \tag{4.1}$$

While matter tells gravity how to behave through the Poisson Equation,

$$\Delta\phi = 4\pi G\rho \quad \text{matter} \implies \text{gravity} \tag{4.2}$$

Here you can see the dual nature between equations (4.1) and (4.2). How can we generalize these equations to the relativistic regime. What we will see is that (4.1) become the geodesic equations, while (4.2) become the Einstein field equations.

Let's find the relativistic version of (4.2).

$$\Delta\phi = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi$$

Here Δ is the Laplacian. The relativistic notation is given by,

$$\Box\phi \equiv \eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}\phi = \left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \Delta\right)\phi \tag{4.3}$$

Where $\Box$ is called the *D'Alembertian*. What about the RHS of (4.2)? $4, \pi, G$ are all constants, so all that remains is ρ. Notice that $\rho \sim M/V$ is a mass over a volume. As we have seen above mass M is like an energy. Furthermore, boosts affect the volume V.

$$p^\mu = \begin{bmatrix} E \\ \vec{p} \end{bmatrix}$$

Now perform a boost on p^μ ($p^\mu \to \tilde{p}^\mu$) using $\Lambda^\alpha{}_\beta$,

$$\Lambda^\alpha{}_\beta = \begin{bmatrix} \gamma & \frac{\gamma v}{c^2} & 0 & 0 \\ \gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.4}$$

As you can see, the energy terms will get mixed up with the momentum terms.

Consider a *perfect fluid* (set of particles) characterized by,

- velocity v
- mass density ρ
- pressure P

A perfect field is called *dust* is pressure $P = 0$. Then $\rho = M/V$ where the total mass M is given by,

$$M = n \cdot m$$

Where n is the number of particles and m is the mass of each particle. Therefore,

$$\rho = \frac{M}{V} = \frac{nm}{V} = \frac{n}{V}m$$

Now the term n/V represents a number density of particles. What is the relativistic version of each of these two terms (n/V and m). Evident m is going to be generalized using the four-momentum p^μ. But what about n/V?

$$N \equiv \frac{n}{V} = \text{ number density}$$

In the co-moving frame (i.e. the frame at which the fluid is at rest) we have, $N = n/V$. Consider a rectangular prism in 3d space with sides $\Delta x, \Delta y, \Delta z$. Then the total volume is $\Delta x \Delta y \Delta z$. Now consider a frame that is not co-moving. Consider this frame moving at velocity $\vec{v} = v_x \hat{x}$. What happens to Δx? Length contraction decreases the width of the box.

$$\Delta x \to \Delta \bar{x} = \frac{1}{\gamma}\Delta x$$

What happens to N? Well since the total number of particles n remains constant, the number density increases,

$$N \to \bar{N} = \frac{n}{\Delta \bar{V}} = \gamma \frac{n}{\Delta x \Delta y \Delta z}$$

Can we see a flux as a number density (i.e. a number of particles crossing an area per unit area of time)? In the co-moving frame, we are moving with the particles so there is no flux. However in the non co-moving frame where $\vec{v} = v_x \hat{x}$, consider a slice along the $\bar{x}$ axis ($\bar{y}\bar{z}$-plane). What is the flux of particles flowing through this slice?

$$\#\text{ particles crossing slice } = \bar{N} \cdot V = \bar{N} \underbrace{\Delta \bar{x}}_{v\Delta \bar{t}} \underbrace{\Delta \bar{y} \Delta \bar{z}}_{\Delta A}$$

The flux then is given by,

$$\text{Flux } = \frac{\#}{\Delta \bar{t} \Delta \bar{A}}$$

Where $\Delta \bar{t} \Delta \bar{A}$ can be thought of as a *spacetime volume* in 3d. So flux is a number density defined with respect to a spacetime 3d volume. Therefore,

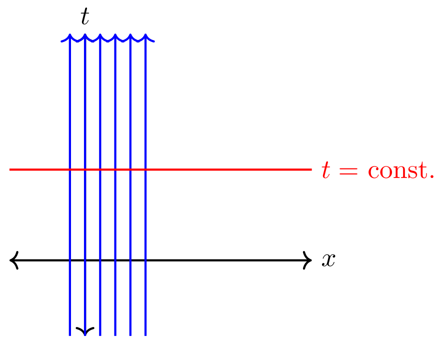
$$\text{Flux } = \bar{N} v = v\gamma N$$

4.3 Duality of Flux and Number Density

In summary of the above considerations,

	Co-moving Frame	Frame (v_x)
# density	$N = \frac{m}{V}$	$\bar{N} = \frac{n}{\bar{V}} = \gamma \frac{n}{V} = \gamma N$
Flux	0	$F_x = \gamma N v_x$

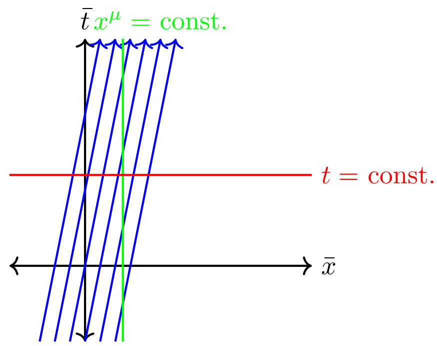
The question becomes, can we see a flux as a number density? In the comoving frame we have:



Thus this represents some sort of flux through a fixed time.

$$\#\text{density} = \frac{\#\text{particles}}{\Delta V} = \frac{\#\text{particles}}{\Delta x \Delta A}$$

Whereas in the non co-moving frame the particles moving at speed v_x:



The relevant notion of relativistic number density is the notion of flux through the $x^\mu = \text{const.}$ We will introduce the "#-flux" vector N^μ. Where at rest,

$$N^\mu = \begin{bmatrix} N \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

And under a boost such as in equation (4.4) where $v = v_x$,

$$N^\mu = \begin{bmatrix} \gamma N \\ \gamma N v_x \\ 0 \\ 0 \end{bmatrix} = N \begin{bmatrix} \gamma \\ \gamma v_x \\ 0 \\ 0 \end{bmatrix} = N v^\mu$$

As a result, the magnitude of N^μ is given by,

$$N^\mu N_\mu = N^2 v^\mu v_\mu = -N^2$$

Now consider the mass density discussed above $\rho = M/V$. At rest we have $\rho = M/V$ with $E = M$ $(c = 1)$. Under a boost with v_x,

$$\begin{bmatrix} E = m \\ 0 \\ 0 \\ 0 \end{bmatrix} \to \begin{bmatrix} \gamma m = \bar{E} \\ m\gamma v \\ 0 \\ 0 \end{bmatrix}$$

Furthermore the volume under the boost becomes $V \to \bar{V} = \frac{1}{\gamma}V$. So $\rho = M/V$ at rest becomes under a boost,

$$\bar{\rho} = \gamma^2 \rho$$

This is similar to the velocity terms but instead we have a γ^2 term instead of γ. Note that density is given by,

$$\rho = \frac{M}{V} = \frac{n}{V} m \tag{4.5}$$

Where m is the mass of a single particle that should be considered the mass given by the four-momentum p^α. The number density should be given by the four-number density N^β. Using this interpretation, (4.5) is expressed under the relativistic regime as,

$$\rho \to N^\alpha \otimes p^\beta = N v^\alpha \otimes m v^\beta = m N v^\alpha \otimes v^\beta = \rho v^\alpha \otimes v^\beta$$

This new quantity $\rho v^\alpha \otimes v^\beta$ will be called $T^{\alpha\beta}$. At rest,

$$v^\alpha = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Which gives $T^{00} = \rho$ and $T^{\alpha\beta} = 0$ whenever $\alpha \neq 0$ or $\alpha \neq 0$. While in frame (v_x),

$$v^\alpha = \begin{bmatrix} \gamma \\ v_x \\ 0 \\ 0 \end{bmatrix}$$

Which gives,

$$T^{00} = \gamma^2 \rho$$

So why does this γ^2 term keep showing up? Since T^{00} is a tensor with two components the γ induced by the boost affects each of the components, namely M and V. The boost affects the tensor for each index,

$$T^{\alpha\beta} \underbrace{\to}_{\text{boost}} \Lambda^\alpha{}_\gamma \Lambda^\beta{}_\delta T^{\gamma\delta} = T^{\bar{\alpha}\beta}$$

This tensor $T^{\alpha\beta}$ is called the *stress energy tensor* and will act as a source of gravity.

- T^{00}: energy density
- T^{0i}: energy flux across i-th surface $(i = 1, 2, 3)$
- T^{i0}: momentum density $(i = 1, 2, 3)$
- T^{ij}: flux of i-th momentum through j-th surface $(i, j = 1, 2, 3)$
  - this is known as the stress tensor that appears when particles have interactions with one another.

4.4 Properties of Stress Energy Tensor

symmetric: $T^{\alpha\beta} = T^{\beta\alpha}$

Here is a "proof" using dimensional arguments.

$$\begin{aligned} T^{0i} &= \text{energy flux} \\ &= \text{density of energy} \times \text{speed of flow} \\ &= \text{density of mass} \times \text{speed of flow} \\ &= \text{density of momentum} \\ &= T^{i0} \end{aligned}$$

conservation: $\frac{\partial}{\partial x^\beta} T^{\alpha\beta} = 0$

To demonstrate this conservation, think of conservation as "what goes in and out of a box encodes the variations of what's inside the box." Consider a cube of side length ℓ aligned to a Cartesian coordinate system. Focusing on energy,

$$\ell^3 \frac{\partial}{\partial t} T^{00} \quad \text{variation of what's inside the box} \tag{4.6}$$

The rate of flow of energy of through each of the 6 faces of the cube is given by,

$$\ell^2 T^{0x}(x=0) - \ell^2 T^{0x}(x=\ell) + \ell^2 T^{0y}(y=0) - \ell^2 T^{0y}(y=\ell) + \ell^2 T^{0z}(z=0) - \ell^2 T^{0z}(z=\ell) \tag{4.7}$$

Therefore by conservation (4.6) must equal (4.7),

$$\ell^3 \frac{\partial}{\partial t} T^{00} = \ell^2 \big(T^{0x}(x=0) - T^{0x}(x=\ell) + T^{0y}(y=0) - T^{0y}(y=\ell) + T^{0z}(z=0) - T^{0z}(z=\ell) \big)$$

Or more cleanly, dividing by ℓ^3 and considering the limit as $\ell \to 0$,

$$\frac{\partial}{\partial t} T^{00} = \lim_{\ell \to 0} \frac{1}{\ell} \sum_{i=1}^{3} \big(T^{0i}(x^i = 0) - T^{0i}(x^i = \ell) \big)$$

Using the definition of partial derivatives,

$$\frac{\partial}{\partial t} T^{00} = -\frac{\partial}{\partial x} T^{0x} - \frac{\partial}{\partial y} T^{0y} - \frac{\partial}{\partial z} T^{0z}$$

Or rearranged,

$$\sum_{\alpha=1}^{3} \frac{\partial}{\partial x^\alpha} T^{0\alpha} = \frac{\partial}{\partial t} T^{0t} + \frac{\partial}{\partial x} T^{0x} + \frac{\partial}{\partial y} T^{0y} + \frac{\partial}{\partial z} T^{0z} = 0$$

By Einstein summation convention,

$$\frac{\partial}{\partial x^\alpha} T^{\beta\alpha} = 0$$

4.5 Early Attempts at Relativistic Gravity

Now what can we propose about gravity? Nordstrom in 1907 proposed to Einstein the idea of taking,

$$\Delta \phi = 4\pi G \rho$$

And replacing this with the D'Alembertian defined as (4.3) and the stress energy tensor,

$$\Box\phi = 4\pi G T^{\alpha}{}_{\alpha}$$

This proposed theory is incorrect because this equation is linear. Here, gravity doesn't *gravitate*.
In 1913, after hard work, Nordstrom proposed,

$$\frac{\Box\phi}{\phi} = 4\pi G T^{\alpha}{}_{\alpha} \tag{4.8}$$

This theory is self consistent by it can be shown that under this theory, light doesn't bend with gravity. This has been proven to be true in our universe so it must also be wrong.

In 1914 however, Einstein and Fohker discovered that (4.8) could be recast as,

$$R = \tilde{T}^{\alpha}{}_{\alpha}$$

Where R is the *Ricci scalar curvature* from the study of Differential Geometry. This new theory is a geometrical theory. This is nice because it is a geometric, scalar, relativistic theory for gravity. Again this theory does not permit for the bending of light just as (4.8); thus it must be wrong. Almost immediately afterwards, Einstein came out with his theory of general relativity in 1915.

4.6 Equivalence Principles

What is the difference between inertial mass and gravitational mass.

- **Inertial mass**: Mass of an object measured by it's resistance to acceleration.
  - An object with high inertial mass with accelerate less than objects with low inertial mass when subject to the same applied force
  - $a = F/m_I$
- **Gravitational mass**: Mass of an object that defines the gravitational force on a system.
  - $m_G\vec{g} = \vec{F}_g$

Note there are other forms of mass like relativistic mass and rest mass.

Experimentally inertial mass m_I and gravitational mass m_G have been found to be identical.

4.6.1 Weak Equivalence Principle

The weak equivalence principle states that test bodies fall with the same acceleration independently of their structure or composition. Formally,

$$\frac{m_I}{m_G} = 1$$

This was experimentally verified by experiments in 1885 by Eötvös. Today we have verified this to 10^{-12}, or 12 decimal places. Exploring this idea,

$$m_I\vec{a} = m_G\vec{g} \implies \vec{a} = \vec{g}$$

Uniform acceleration cannot be distinguished from a uniform gravitational field.

Consider the ISS orbiting at an altitude of $h = 400\,\text{km}$. The gravitational field is roughly 90% that of the surface gravitational field.

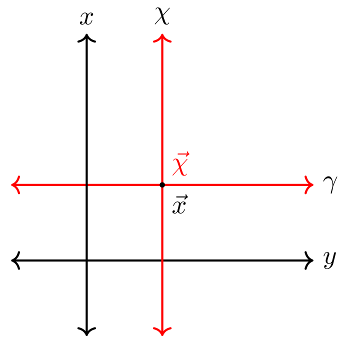
$$g = \frac{GM}{R+h}$$

Consider a particle in space characterized by position $\vec{x}$.

$$m_I \frac{\mathrm{d}^2 \vec{x}}{\mathrm{d}t^2} = m_G \vec{g}$$

Now consider a new frame $\vec{\chi}$ such that $\frac{\mathrm{d}^2 \vec{\chi}}{\mathrm{d}t^2} = \vec{a}$ with coordinates in the old frame given by,

$$\vec{x}\,' = \vec{x} = \vec{\chi}$$



Therefore,

$$m_I \frac{\mathrm{d}^2 \vec{x}\,'}{\mathrm{d}t^2} = m_I \frac{\mathrm{d}^2 \vec{x}}{\mathrm{d}t^2} - m_I \frac{\mathrm{d}^2 \vec{\chi}}{\mathrm{d}t^2} = m_G \vec{g} - m_I \vec{a} = m_G(\vec{g} - \vec{a})$$

Which confirms again that uniform acceleration is indistinguishable from a uniform gravitational field. Also note that the first falling frame is an inertial frame.

$$m_I \frac{\mathrm{d}^2 \vec{x}\,'}{\mathrm{d}t^2} = m_G(\vec{g} - \vec{g}) = \vec{0}$$

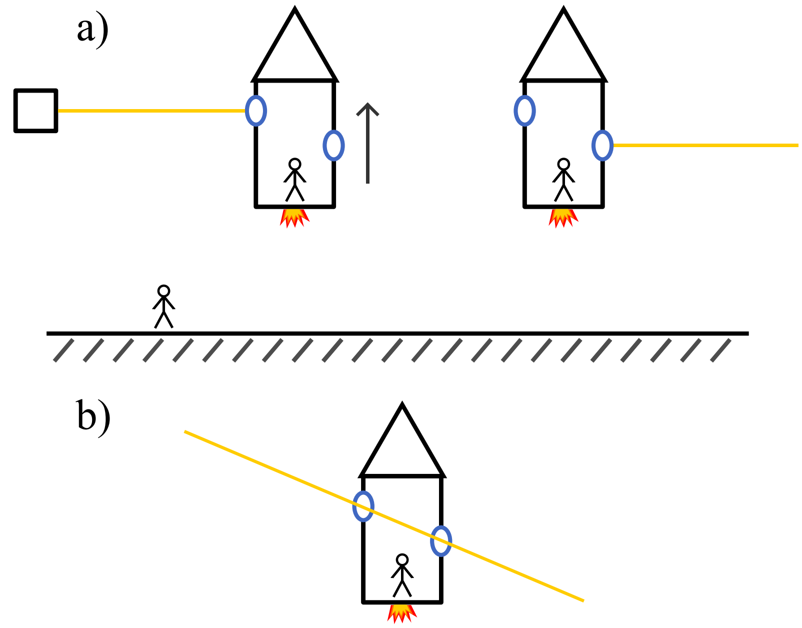
4.6.2 Tidal Forces

Consider a spaceship large enough to experience the change in the gravitational field from one end to another. This force gradient causes stress and strain on the spaceship. These forces are called Tidal forces. To avoid tidal forces, only consider objects in a *local inertial frame* where the field is *uniform*.

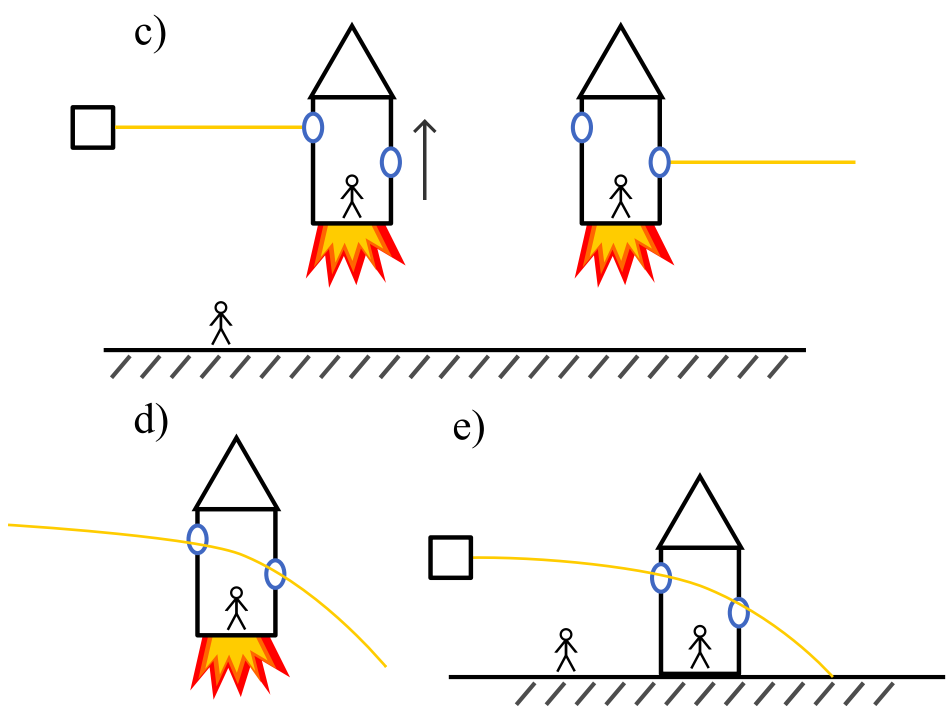
4.6.3 Einstein's Strong Equivalence Principle

Free-falling in local inertial frame (where $\vec{g}$ is uniform), the results of all experiments will be indistinguishable from the results of the same experiments performed in a inertial frame in Minkowski spacetime.

Consider the case of a rocket with initially no acceleration $\vec{a} = \vec{0}$ and constant velocity $\vec{v} = \text{const}$. Another ship comes along and shoots a laser. Suppose there is a window on the left w_L and a window on the right w_R. Now suppose the spaceship gets lucky and the laser passes through w_L and then w_R (a). From the perspective of the captain of the ship, the light moves as a straight line (b). Nothing too complicated.



Now consider the rocket moving with constant acceleration $\vec{a} = \text{const}$ (c). Additionally, the fired laser luckily misses the ship and passed through w_L and then w_R. From the perspective of the captain, the light travels along a curved trajectory (d). Nothing too complicated. However by the Einstein's Strong Equivalence principle, this experiment must be identical to a stationary spaceship in a uniform gravitational field. Therefore, in a uniform gravitational field with everything else stationary, light **must** travel along a curved trajectory (e). Gravity bends light!



We must be able to measure this effect. If the light falls in the gravitational field, it must gain energy $E = \hbar\omega$ and thus it will be blue shifted. Similarly, if the light climbs out of the gravitational field, it will lose energy and be red shifted. To explore this idea, consider a rocket with two individuals, one at the top of the rocket (A) and one at the bottom (B) (separated by height h). The rocket will be accelerating with $\vec{g}$ in just the $\hat{z}$ direction which gives,

$$\ddot{z}_A = \ddot{z}_B = g$$

Therefore,

$$z_A(t) = h + \frac{1}{2}gt^2$$

$$z_B(t) = \frac{1}{2}gt^2$$

At $t = 0$, A sends a light signal to B and is received by B at time $t_1 > t_0 = 0$.

$$ct_1 = z_A(0) - z_B(t_1)$$

At time $t = \Delta\tau_A$, A emits another light signal to B and it will be received at time $t_1 + \Delta\tau_B$,

$$c(t_1 + \Delta\tau_B - \Delta\tau_A) = z_A(\Delta\tau_A) - z_B(t_1 + \Delta\tau_B)$$

With these two relations, we want to get an expression that relates $\Delta\tau_A$ to $\Delta\tau_B$. Considering a non-relativistic regime $c \gg 1$, and at 1st order in $\Delta\tau_i$ with $i = A, B$,

$$\Delta\tau_B = \Delta\tau_A\left(1 - \frac{gh}{c^2}\right)$$

This is strange. If there were no acceleration $\Delta\tau_B = \Delta\tau_A$. However, there is an extra factor of due to the acceleration $\vec{g}$. Because of the acceleration, the second beam of light had to travel a farther distance making $\Delta\tau_B < \Delta\tau_A$. If $\Delta\tau$ is taken to be the period of light waves,

$$\Delta\tau = \frac{\lambda}{c}$$

Thus,

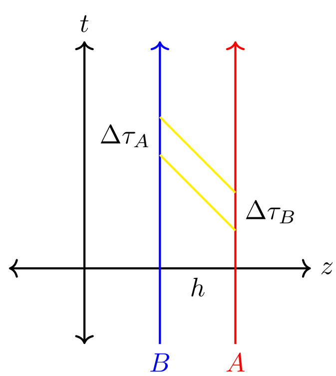
$$\lambda_B = \lambda_A\left(1 - \frac{gh}{c^2}\right)$$

Which gives use a similar relation between λ_A and λ_B, namely $\lambda_B < \lambda_A$. This corresponds to a **blue shift**. Following a similar argument with $g \to -g$ it becomes that $\lambda_B > \lambda_A$ which corresponds to a **red shift**.

However, by Einstein's Strong Equivalence Principle, the *exact* same effect should occur while stationary under a uniform gravitational field $\vec{g}$. We can perform this experiment on earth and measure the effect. This was known as the *Pound Rebka experiment*. This effect is very small. On earth,

$$\frac{gh}{c^2} \sim 10^{-15}$$

What does this experiment look like as a spacetime diagram,



There is a problem here. Since the field $\vec{g}$ is uniform, it must be that the two light rays are *parallel*. Since the gravitational potential at A and B is given by ϕ_A, ϕ_B and they are equal,

$$\Delta\tau_B \approx \left(1 - \frac{\phi_A - \phi_B}{c^2}\right)\Delta\tau_A = \Delta\tau_A \tag{4.9}$$

However, we have just shown that,

$$\Delta\tau_B < \Delta\tau_A$$

Which indicates that these lines are converging and thus should **not** be parallel. Nonetheless, this problem can be resolved when considering geometries like the surface of a sphere where *great circles* are both parallel and intersecting. This convinced Einstein that he would have to learn *differential geometry*.

Before moving to differential geometry, let us recall that for the Minkowski metric,

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^2$$

Now consider a modified metric,

$$\mathrm{d}\tau^2 = (1 + 2\phi(\vec{x}))\mathrm{d}t^2 - (1 - 2\phi(\vec{x}))\mathrm{d}\vec{x}^2 \tag{4.10}$$

This is known as the *static weak field metric* where ϕ is the gravitational potential. Can this metric recover some of the issues discussed above? If A and B are at rest $\Delta\vec{x}^2 = 0$,

$$\Delta\tau_A^2 = (1 + 2\phi(x_A))\Delta t^2$$

$$\Delta\tau_B^2 = (1 + 2\phi(x_B))\Delta t^2$$

Which gives,

$$\Delta\tau_A^2 = \frac{1 + 2\phi(x_A)}{1 + 2\phi(x_B)}\Delta\tau_B^2$$

Where a Taylor series expansion gives,

$$\Delta\tau_A \approx (1 + \phi(x_A) - \phi(x_B))\Delta\tau_B$$

Which can be seen to mimic (4.9). This motivates how gravity can be described as a modified metric which in turn implies a gravity affects the space in which the system lives.

5 Differential Geometry

We known the flat $\mathbb{R}^n$ space very well. How can be relate different geometries back to $\mathbb{R}^n$. Consider the surface of the earth. How can we measure *where we are*? Consider someone living in Waterloo, they might have a map that has coordinates describing where they are. Similarly for someone in Toronto with a different map, they would have coordinates that describe where they are. In order for these two people to communicate, there needs to be a translation between the coordinates on one map to another. This is the spirit of differential geometry. By stitching together maps (charts), we can get a description of geometry through the non-trivial overlaps between adjacent maps. At any given space in our manifold, we can define a set of tangent vectors which span a *tangent bundle* or tangent plane. These tangent planes can form a vector space throughout the geometry. Through these spaces, we can define co-vectors and thus tensors like metrics without these spaces. We will also explore the idea of a derivative of a tensor that relates difference tangent planes.

5.1 Definitions

Open Ball in $\mathbb{R}^n$: An open ball in $\mathbb{R}^n$ centered at y is the set of all $x \in \mathbb{R}^n$ such that,

$$\left\{ x \in \mathbb{R}^3 \mid \text{for a given } y \in \mathbb{R}^n, |x-y| < R \text{ with } |x-y|^2 = \sum_i (x-y)_i^2 \right\}$$

Where $R \in \mathbb{R}_{>0}$.

Open set: An open set in $\mathbb{R}^n$ is built from the union of open balls.
Chart: Let M be a set not necessarily $\mathbb{R}^n$. A **chart** or coordinate system $(\phi, \mathcal{U})$ consists of a subset $\mathcal{U}$ of M with a one-to-one map,

$$\phi : \mathcal{U} \to \mathbb{R}^n$$

Such that $\phi(\mathcal{U}) = V$ is an open set (in $\mathbb{R}^n$). We can say that $\mathcal{U} = \phi^{-1}(V)$. We can use V to induce the topology of our topological space. By notation, we usually note $\phi(p)$ where $p \in \mathcal{U}$, as,

$$\phi(p) \equiv x^\mu(p) \quad \phi \text{ acts as the } \textit{notion} \text{ or coordinates in } \mathbb{R}^n \text{ at } p$$

Atlas: An atlas is a collection of charts $(\mathcal{U}_\alpha, \phi_\alpha)$ such that the union of all $\mathcal{U}_\alpha$ covers all of M,

$$\cup_\alpha \mathcal{U}_\alpha = M$$

If two charts happen to overlap, (i.e. $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$), then the map $\phi_\alpha \circ \phi_\beta^{-1}$ takes points in $\phi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$ to $\phi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$. We say that the chart is C^∞ (or smooth) if ϕ_α is C^∞ (infinitely differentiable). If all the charts are C^∞ then $\phi_\alpha \circ \phi_\beta^{-1}$ and its inverse are C^∞.

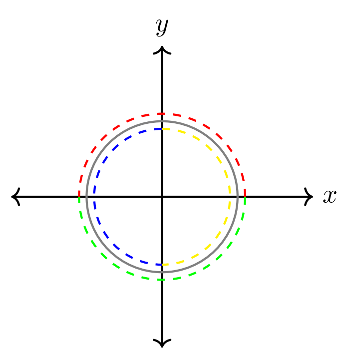
Manifold: A manifold is a set M equipped with a maximal atlas (i.e. the one that contains all the possible charts). If the atlas is defined by $\phi_\alpha : M \to \mathbb{R}^n, \forall \alpha$ then M is a manifold of dimension n.

5.1.1 Examples of Manifolds

Consider S_1 the circle in the plane $\mathbb{R}^2$,

$$S^1 = \left\{ (x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \right\}$$

Now consider the open sets consisting of $y > 0$ (red) and $y < 0$ (green), as well as $x > 0$ (yellow) and $x < 0$ (blue).



The four points of the form, $(x,y) = (1,0), (-1,0), (0,1), (0,-1)$, all belong to a single open set on S^1. However, for regions that belong to two separate open sets, we need a map that takes us from one open set to the next. For example, to map from yellow to red,

$$y_{\text{red}} = \sqrt{1 - x_{\text{yellow}}^2} \tag{5.1}$$

This acts as the form $\phi_x \circ \phi_y^{-1}(x)$. Is this composite map C^∞? Note that (5.1) is only not differentiable at $x = 0$. However, this point doesn't not belong the intersection of the red and yellow regions. Therefore, (5.1) is C^∞. More more, see Caroll's book (online) for differential geometry.

5.2 Tangent Vectors

Before we defined the notion of tangent vectors, we will need to first discuss trajectories.

5.2.1 Trajectories

Definition: A C^∞ curve in a manifold M is a map $\mathcal{C} : \mathbb{R} \supset I \to M, \tau \to \mathcal{C}(\tau)$ were I is an interval on $\mathbb{R}$ such that,

$$\forall \alpha : \phi_\alpha \circ \mathcal{C} \text{ is a } C^\infty \text{ map}$$

For points p in M get mapped to $x^\mu(P)$ in $\mathbb{R}^m$ through ϕ_m along the interval I parametrized by τ. In summary, we take an interval I in $\mathbb{R}$ and then map it to a trajectory in M using $\mathcal{C}$ which is taken locally to $\mathbb{R}^n$ by ϕ_α.

$$\mathcal{C} : I \to M$$

$$\phi_\alpha \circ \mathcal{C}(\tau) \equiv x^\mu(\tau)$$

Definition: We note $\mathcal{C}^\infty(M)$ the **algebra** of C^∞ functions of the form $f : M \to \mathbb{R}$, with point-wise product given by,

$$(f_1 \cdot f_2)(p) = f_1(p) f_2(p)$$

Where $P \in M$.
Definition: Let $\mathcal{C} : I \to M$ be a smooth curve with $\mathcal{C}(0) = p$. The **tangent vector** to $\mathcal{C}$ at p is the linear map $X_p : \mathcal{C}^\infty(M) \to R$ given by,

$$X_p(f) = \frac{\mathrm{d}}{\mathrm{d}\tau} f(\mathcal{C}(\tau))|_{\tau=0}$$

Notation:

$$f \circ \mathcal{C} = f \circ \phi_\alpha^{-1} \circ \phi_\alpha \circ \mathcal{C}$$

Using the combination $F = f \circ \phi_\alpha^{-1}$ which is a map $F : \mathbb{R}^n \to \mathbb{R}$ (since $\phi_\alpha \circ \mathcal{C} : I \to R^n$) which gives a more convenient expression,

$$f \circ \mathcal{C}(\tau) = F(x^\mu(\tau))$$

Using the coordinate system, this becomes,

$$\begin{aligned} X_P(f) &= \frac{\mathrm{d}}{\mathrm{d}\tau}(F(x^\mu(\tau))|_{\tau=0} \\ &= \underbrace{\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right|_{\tau=0}}_{\text{components}} \underbrace{\left(\frac{\partial}{\partial x^\mu} F(x^\mu)\right|_p}_{\text{basis}} \quad \text{Chain rule.} \end{aligned}$$

Exploring this notation motivates the separation of the coordinates and basis of the tangent vector $X_p(f)$. Focusing on the specific curves $\mathcal{C}_k(\tau)$ such that,

$$x^\mu(\mathcal{C}_k(\tau)) = \left(x^0(p), \ldots, x^k(p) + \tau, \ldots, x^n(n)\right)$$

This addition of τ along each direction allows for the "exploration" of the k-th direction by an amount τ.

$$\mathcal{C}_1 : (x(p) + \tau, y(p))$$

$$\mathcal{C}_2 : (x(p), y(p) + \tau)$$

This gives the form,

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}(\mathcal{C}_k(\tau))|_{\tau=0} = \delta^\mu{}_k$$

So that the tangent vector of $\mathcal{C}_k$ is $\frac{\partial}{\partial x^k}|_p$.
Definition: The set of all tangent vectors at a point p is the tangent plane T_pM at p.

Theorem: If M has dimension n then T_pM has dimension n, making $T_pM \sim \mathbb{R}^n$. The basis of T_pM is given by $\frac{\partial}{\partial x^\mu}|_p$.

Comments:

- In the tangent space T_pM, a general vector is $X = X^\mu \frac{\partial}{\partial x^\mu} = X^\mu \partial_\mu$ (noting, down in the ∂ term means up in the denominator by notation).
- The basis $\frac{\partial}{\partial x^\mu}$ is coordinate dependent. Another coordinate system could be $\frac{\partial}{\partial y^\mu}|_p$.

These different basis are of course related by the Jacobian (Wald p.17),

$$\left.\frac{\partial}{\partial x^\mu}\right|_p = \left.\frac{\partial y^\alpha}{\partial x^\mu}\right|_p \left.\frac{\partial}{\partial y^\alpha}\right|_p$$

So a generic vector in T_pM is $X = X^\mu \frac{\partial}{\partial x^\mu} = \tilde{X}^\mu \frac{\partial}{\partial y^\mu}$. How do the components X^μ relate to $\tilde{X}^\mu$?

$$X^\mu \frac{\partial}{\partial x^\mu} = \tilde{X}^\alpha \frac{\partial}{\partial y^\alpha} = X^\mu \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial}{\partial y^\alpha}$$

This gives the relation,

$$X^\mu \to \tilde{X}^\mu = X^\alpha \frac{\partial y^\mu}{\partial x^\alpha}$$

Definition: We define a cotangent space T_p^*M as the set of linear maps $T_pM \to \mathbb{R}$. It has dimension n. We note the basis of T_p^*M by $\mathrm{d}x^\mu$ (sometimes called a form).

$$\mathrm{d}x^\mu(\partial_\alpha) = \delta^\mu{}_\alpha \quad \text{analogously to } f^\mu(e_\alpha) = \delta^\mu{}_\alpha$$

Remark: Both the tangent space T_pM and cotangent space T_p^*M are in general not contained in the manifold M. Only in flat Minkowski space is this is case.

5.3 Tensor Calculus

Definition: A tensor of type (r, q) on T_pM is a multi-linear map:

$$T : \underbrace{T_p^*M \times \cdots \times T_p^*M}_{r} \times \underbrace{T_pM \times \cdots \times T_pM}_{q} \to \mathbb{R}$$

Examples:

Tensor	Rank
X^μ	$(1, 0)$
X_μ	$(0, 1)$
$g_{\alpha\beta}$	$(0, 2)$

Definition: A type $(0, 2)$ tensor $g_{\alpha\beta}$ which is non-degenerate ($v^\mu g_{\mu\alpha} w^\alpha = 0 \quad \forall w^\alpha \implies v^\mu = 0$) and symmetric ($g_{\alpha\beta} = g_{\beta\alpha}$) is a metric tensor. The notion of signature defined in $\mathbb{R}^n$ is the same as before.
Definition: We consider the union of the tangent spaces T_pM for all $p \in M$ (respectively T_p^*M). We note,

$$T(M) = \bigcup_{p\in M} T_pM \qquad T^*(M) = \bigcup_{p\in M} T_p^*M$$

A **vector field** is a map $X : M \to T(M), p \to X_p$. Likewise a **tensor field** is,

$$T : M \to T^*(M) \times \cdots \times T^*(M) \times T(M) \cdots \times T(M)$$

Or for a specific point p, $p \to T_p$. From now on we will work with tensor fields. We extend all the formalism done in the Minkowski space.

- Metric gives us the norm of vectors $v^\alpha g_{\alpha\beta} v^\beta = L$ where $v^\alpha, g_{\alpha\beta}$ are associated with a tangent plane at a point $T_p M$
- The sign of L determines the type of vector
  - $L > 0 \to V^\alpha$ is space-like
  - $L = 0 \to V^\alpha$ is light-like
  - $L < 0 \to V^\alpha$ is time-like
- A curve in M has type that follows from the tangent vector v^α
  - M is space-like if tangent vector is space-like
  - M is light-like if tangent vector is light-like
  - M is time-like if tangent vector is time-like
- The arclength of a curve is given by,

$$\tau = \int \sqrt{\left| \frac{\mathrm{d}x^\mu}{\mathrm{d}\ell} g_{\mu\alpha} \frac{\mathrm{d}x^\alpha}{\mathrm{d}\ell} \right|}$$

Where ℓ is a curvilinear parameter of the curve and $\frac{\mathrm{d}x^\mu}{\mathrm{d}\ell}$ is the tangent vector. The line element is given by,

$$\mathrm{d}\tau^2 = |g_{\alpha\mu} \mathrm{d}x^\mu \mathrm{d}x^\alpha|$$

We will physically interpret the arclength τ to be proper time.

- $\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ = tangent vector = relativistic velocity
- $\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}$ = relativistic acceleration

How can we generalize Newton's law $\vec{F}_g = m_I \vec{a}$. Using the scalar gravitational field,

$$-m_G \vec{\nabla}\phi = m_I \vec{a}$$

Since the weak equivalence principle gives us $m_G = m_I$. Therefore,

$$-\vec{\nabla}\phi = \vec{a}$$

What is the relativistic version? Like us assume we are in a *free falling* frame given by coordinate ξ^α. Then according to Einstein's strong equivalence principle, the acceleration is just zero,

$$a^\mu = \frac{\mathrm{d}^2\xi^\mu}{\mathrm{d}\tau^2} = 0$$

Since we are in a free falling frame, physics must behave like the physics in Minkowski spacetime. The line element is given by,

$$\mathrm{d}\tau^2 = \eta_{\mu\alpha} \mathrm{d}\xi^\mu \mathrm{d}\xi^\alpha$$

Where $\eta_{\mu\alpha}$ is the Minkowski metric. Now let us perform a change of coordinates and move *away* from the free falling frame.

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

This transformation is transformed by the Jacobian. However notice that,

$$\frac{\mathrm{d}^2\xi^\mu}{\mathrm{d}\tau^2} = \frac{\mathrm{d}v^\mu}{\mathrm{d}\tau}$$

But as shown on the assignment, v^μ's coordinates do not transform well. There are extra terms generated.

$$\frac{\partial \xi^\mu}{\partial x^\alpha} \quad \text{Jacobian}$$

$$\frac{\partial x^\alpha}{\partial \xi^\mu} \quad \text{Inverse Jacobian}$$

Which gives via the inverse relation,

$$\frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial \xi^\rho} = \delta^\mu{}_\rho$$

Which gives,

$$\begin{aligned} \mathrm{d}\tau^2 &= \eta_{\mu\alpha}\mathrm{d}\xi^\mu \mathrm{d}\xi^\alpha \\ &= \eta_{\mu\alpha}\frac{\partial \xi^\mu}{\partial x^\gamma}\frac{\partial \xi^\mu}{\partial x^\beta}\mathrm{d}x^\gamma \mathrm{d}x^\beta \\ &= g_{\gamma\beta}\mathrm{d}x^\gamma \mathrm{d}x^\beta \end{aligned}$$

But if we have the velocity to be,

$$v^\mu = \frac{\mathrm{d}\xi^\mu}{\mathrm{d}\tau} = \frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}$$

This implies acceleration is given by *product rule*,

$$\begin{aligned} 0 &= a^\mu \\ &= \frac{\mathrm{d}}{\mathrm{d}\tau}v^\mu \\ &= \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\right) \\ &= \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial \xi^\mu}{\partial x^\alpha}\right) + \frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\right) \quad \text{Product rule.} \\ &= \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\partial^2 \xi^\mu}{\partial x^\alpha \partial x^\beta}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} + \frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\tau^2} \quad \text{Chain rule.} \end{aligned}$$

Multiply by the inverse Jacobian on both sides (and send $\alpha \to \gamma$ for ease of notation),

$$0 = \frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\tau^2} + \underbrace{\frac{\partial x^\alpha}{\partial \xi^\mu}\frac{\partial^2 \xi^\mu}{\partial x^\gamma \partial x^\beta}}_{\Gamma^\alpha{}_{\gamma\beta}}\frac{\mathrm{d}x^\gamma}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau}$$

Where $\Gamma^\alpha{}_{\gamma\beta}$ are called the Christoffel symbols. Notice that they are symmetric in the lower indices,

$$\Gamma^\alpha{}_{\gamma\beta} = \Gamma^\alpha{}_{\beta\gamma}$$

There is relationship between $\Gamma^\alpha{}_{\beta\gamma}$ and the metric. To illustrate this, examine that the metric transforms well,

$$g_{\alpha\beta} = \eta_{\mu\nu}\frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\partial \xi^\nu}{\partial x^\beta}$$

This relationship can then be derived to be,

$$\Gamma^\alpha{}_{\gamma\beta} = \frac{1}{2}g^{\alpha\lambda}\left(\frac{\partial g_{\lambda\gamma}}{\partial x^\mu} + \frac{\partial g_{\mu\lambda}}{\partial x^\gamma} - \frac{\partial g_{\mu\gamma}}{\partial x^\lambda}\right)$$

This implies,

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\tau^2} + \Gamma^\alpha{}_{\gamma\beta}\frac{\mathrm{d}x^\gamma}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau}$$

Which implies the relationship between acceleration and the change of coordinates characterized by ${\Gamma^\alpha}_{\gamma\beta}$,

$$a^\alpha = -{\Gamma^\alpha}_{\gamma\beta} v^\gamma v^\beta$$

Can we retrieve in a non relativistic limit $\vec{a} = -\vec{\nabla}\phi$? To do this we need to convert show the analogy of this geometrical equation,

$$\tilde{a}^\mu = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + {\Gamma^\mu}_{\alpha\beta} v^\alpha v^\beta = 0 \tag{5.2}$$

To the gravitational force equation,

$$\vec{a} = -\vec{\nabla}\phi \tag{5.3}$$

In order to reveal this connection, we will perform a non-relativistic limit to (5.2) and a weak field stationary metric to (5.3). Using $c = 1$ and $|\vec{v}| \ll 1$ we have,

$$v^\alpha = \begin{bmatrix} \gamma \\ \gamma v^1 \\ \gamma v^2 \\ \gamma v^3 \end{bmatrix}$$

Making for $i = 1, 2, 3$, the real space velocity components v^i/v^0. Therefore we have $|\vec{v}/v_0| \ll 1$. The non-relativistic limit becomes,

$$|\vec{v}| \ll v_0$$

So therefore the time components dominate,

$${\Gamma^\alpha}_{\gamma\beta} v^\gamma v^\beta \approx {\Gamma^\alpha}_{00} v^0 v^0$$

By definition of ${\Gamma^\alpha}_{\alpha\beta}$,

$${\Gamma^\alpha}_{00} = -\frac{1}{2} g^{\mu\alpha} \frac{\partial g_{00}}{\partial x^\mu} \tag{5.4}$$

Now we will perform a weak field stationary metric on (5.3). Therefore for a general metric $g_{\alpha\beta}$ we can treat it as a Minkowski metric plus a tiny perturbation $|h_{\alpha\beta}| \ll 1$,

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \tag{5.5}$$

As such we will work in a 1st order approximation in $h_{\alpha\beta}$. What is $g^{\alpha\beta}$?

$$g^{\alpha\beta} g_{\beta\gamma} = {\delta^\alpha}_\gamma$$

Subbing in (5.5) the LHS becomes,

$$\left(\eta^{\alpha\beta} - h^{\alpha\beta}\right)\left(\eta_{\beta\gamma} + h_{\beta\gamma}\right) = \eta^{\alpha\beta}\eta_{\beta\gamma} - h^{\alpha\beta}\eta_{\beta\gamma} + \eta^{\alpha\beta} h_{\beta\gamma} - h^{\alpha\beta} h_{\beta\gamma}$$

Noting that $\eta^{\alpha\beta} - h^{\alpha\beta}$ is a candidate for the inverse metric which will be justified after the fact for a first order approximation. $h^{\alpha\beta} h_{\beta\gamma}$ is second order in $h_{\alpha\beta}$ so it is negligible. Also, $\eta^{\alpha\beta}\eta_{\beta\gamma} = {\delta^\alpha}_\gamma$ by inverse property and the remaining terms cancel out since $h_{\alpha\beta}$ must be symmetric (because it's a metric). Therefore out inverse candidate is given by the assumed,

$$g^{\alpha\beta} = \eta^{\alpha\beta} - h^{\alpha\beta}$$

Then using (5.4)

$$\begin{aligned}
{\Gamma^\mu}_{00} &= -\frac{1}{2} g^{\mu\alpha} \frac{\partial g_{00}}{\partial x^\alpha} \\
&= -\frac{1}{2}\left(\eta^{\mu\alpha} - h^{\mu\alpha}\right) \frac{\partial}{\partial x^\alpha}\left(\eta^{00} + h^{00}\right) \\
&= -\frac{1}{2}\left(\eta^{\mu\alpha} - h^{\mu\alpha}\right) \frac{\partial}{\partial x^\alpha}\left(h^{00}\right) \quad \text{Since } \eta^{00} = -1 \text{ in Cartesian coords}
\end{aligned}$$

$$= -\frac{1}{2}(\eta^{\mu\alpha})\frac{\partial}{\partial x^\alpha}\left(h^{00}\right) \quad \text{Ignoring terms that are second order in } h_{\alpha\beta}$$

Recall when discussing the bending of light, (4.10) reproduced the right results. We should expect the following metric,

$$g_{\alpha\beta} = \eta_{\alpha\beta} - 2\delta_{\alpha\beta}\phi(\vec{x})$$

to hold. This motivates $h_{\alpha\beta}$ to take the form $2\delta_{\alpha\beta}\phi(\vec{x})$. Note that this metric is *stationary*. Meaning that,

$$\partial_t h_{\alpha\beta} = \partial_t 2\delta_{\alpha\beta}\phi(\vec{x}) = 0$$

Therefore,

$$\tilde{a}^\mu = \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = -\gamma^\mu{}_{\alpha\beta}v^\alpha v^\beta \approx -\gamma^\mu{}_{00}v^0 v^0$$

Becomes,

$$\tilde{a}^\mu = +\frac{1}{2}\eta^{\mu\alpha}\left(\frac{\partial}{\partial x^\alpha}h_{00}\right)v^0 v^0$$

Which has time component $\mu = 0$,

$$\tilde{a}^0 = +\frac{1}{2}\eta^{0\alpha}\left(\frac{\partial}{\partial x^\alpha}h_{00}\right)v^0 v^0 = 0$$

Since $h_{\alpha\beta}$ is stationary. It also has space components $i = 1, 2, 3$,

$$\tilde{a}^i = \frac{\mathrm{d}^2 x^i}{\mathrm{d}\tau^2} = +\frac{1}{2}\eta^{i\alpha}\left(\frac{\partial}{\partial x^\alpha}h_{00}\right)v^0 v^0$$

Noting that the v^0 terms are (again with $c = 1$),

$$v^0 = \frac{\mathrm{d}x^0}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}$$

Therefore,

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}\tau^2} = +\frac{1}{2}\eta^{i\alpha}\left(\frac{\partial}{\partial x^\alpha}h_{00}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2$$

In the non-relativistic limit $\tau \to t$. Therefore,

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = +\frac{1}{2}\eta^{i\alpha}\left(\frac{\partial}{\partial x^\alpha}h_{00}\right)$$

Noting by motivation above that $h_{00} = -2\phi(\vec{x})$,

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = +\frac{1}{2}\eta^{i\alpha}\left(\frac{\partial}{\partial x^\alpha}(-2\phi(\vec{x}))\right)$$

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = -\eta^{i\alpha}\left(\partial^i\phi\right)$$

Which by sifting of $\alpha = i, \eta^{ii} = 1$,

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = -\partial^i\phi$$

$$\vec{\tilde{a}} = -\vec{\nabla}\phi$$

Therefore Gravity is not a force by a curvature or geometry of spacetime.

5.4 Differential Geometry Summary

In summary, our aim is to take an interval $I \subset \mathbb{R}$ parametrized by curvilinear parameter τ and embed it into the manifold M. In doing so, we ascribe for every τ a point $\mathcal{C}(\tau)$ in the manifold. This creates a curve in M.Since dealing directly with abstract points p in the manifold is difficult, we will use charts to assign coordinates to the points $\mathcal{C}(\tau)$ by using ϕ which gives $x^\mu(\mathcal{C}(\tau))$. Now to discuss smoothness of this curve, we need to use $F : f \circ \phi^{-1} : \mathbb{R}^m \to \mathbb{R}$. Moreover, we can define an $X : \mathcal{C}^\infty(M) \to \mathbb{R}$ to map from f (which makes points in M to $\mathbb{R}$ regardless of the dimension of M) to real numbers $\mathbb{R}$. Then,

$$X(f) = \left.\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right|_{\tau=0} \left.\frac{\partial F}{\partial x^\mu}\right|_{\phi(p)}$$

This motivates the convenient tensor notation for X,

$$X = X^\mu \partial_\mu$$

5.5 Covariant Derivative

Let us continue to explore,

$$m\vec{a} = \vec{F}_g$$

And manipulate the derived expression,

$$m\big(a^\mu + \Gamma^\mu{}_{\alpha\beta} v^\alpha v^\beta\big) = 0$$

Using the definition of a^μ,

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\partial}{\partial x^\alpha}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = v^\alpha \frac{\partial}{\partial x^\alpha} v^\mu$$

Therefore,

$$v^\alpha \frac{\partial}{\partial x^\alpha} v^\mu + \Gamma^\mu{}_{\alpha\beta} v^\alpha v^\beta = 0$$

Which suggests,

$$v^\alpha \left(\frac{\partial}{\partial x^\alpha} v^\mu + \Gamma^\mu{}_{\alpha\beta} v^\beta\right) = 0$$

Where the **covariant derivative** is given by, $\nabla_\alpha v^\mu$ where,

$$\nabla_\alpha v^\mu = \frac{\partial}{\partial x^\alpha} v^\mu + \Gamma^\mu{}_{\alpha\beta} v^\beta$$

This leads to something known as the **geodesic equation**,

$$v^\alpha \nabla_\alpha v^\mu = 0$$

With v^μ being the tangent vector to the geodesic. More formally, the covariant derivate is the *notion of derivative that transforms well under a change of coordinates.*

Definition: The *covariant derivate* ∇ on a manifold M is a map $\nabla : T(M) \times T(M) \to T(M)$ where $T(M) = \cup_{p\in M} T_p M$. It takes two vector fields X, Y and maps it to $\nabla_X Y$. It does so such that the following properties hold,

$$\nabla_{fX+gZ} Y = f\nabla_X Y + g\nabla_Z Y \quad \text{Linearity} \tag{5.6}$$

$$\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z \quad \text{Linearity} \tag{5.7}$$

$$\nabla_X (fY) = (\nabla_X f)Y + f(\nabla_X Y) \quad \text{Product rule}$$

Where $X, Y : \mathcal{C}^\infty(M) \to \mathbb{R}$,

$$f \to X(f) = X^\mu \partial_\mu f$$

Where f comes from $F = f \circ \phi^{-1}$. The covariant derivative on a scalar $\nabla_X f$ is given by,

$$\nabla_X f \equiv X(f) = X^\mu \frac{\partial}{\partial x^\mu} f$$

Consider a generic $\nabla_X Y$ and take X in the basis $X^\mu = \delta^{\alpha\mu}$ so that,

$$X = X^\mu \partial_\mu = \delta^{\alpha\mu}\partial_\mu = \partial_\alpha$$

Which suggests the notation for this particular basis,

$$\nabla_\alpha Y \equiv \nabla_{e_\alpha} Y$$

What is $\nabla_X Y$ in terms of the basis e_α of $T_p(M)$? Let us examine the action of the covariant derivative directly on the basis. Since this quantity $\nabla_\alpha e_\beta$ is an element in the tangent space, it should be able to be written with respect to the basis e_α,

$$\nabla_\alpha e_\beta \equiv \Gamma^\gamma{}_{\alpha\beta} e_\gamma$$

Where $\Gamma^\gamma{}_{\alpha\beta}$ is the coefficients of the *connection* in the basis e_α. The *Christoffel symbols* are an example. Therefore what is the expression for $\nabla_X Y$? Writing X, Y as their basis expansions (noting that $Y = Y^\alpha e_\alpha$ where Y^α are each functions because Y is a vector field and e_α are just basis vectors),

$$\nabla_X Y = \nabla_{X^\alpha e_\alpha}\left(Y^\beta e_\beta\right)$$

Where X^α, Y^β are components of the vector field at a point in the manifold with respect to the corresponding basis $e_\alpha \in T_p(M)$. They are functions of p the point in the manifold. Using (5.6),

$$\nabla_X Y = X^\alpha \nabla_{e_\alpha}(Y) \tag{5.8}$$

Now using (5.7) to expand out Y into it's components,

$$\begin{aligned}
\nabla_X Y &= \nabla_X\left(Y^\beta e_\beta\right) \\
&= \nabla_X\left(Y^0 e_0 + Y^1 e_1 + \cdots\right) \\
&= \nabla_X\left(Y^0 e_0\right) + \nabla_X\left(Y^1 e_1\right) + \cdots \\
&= X\left(Y^0\right)e_0 + Y^0 \nabla_X e_0 + X\left(Y^1\right)e_1 + Y^1\nabla_X e_1 + \cdots \\
&= X\left(Y^\beta\right)e_\beta + Y^\beta \nabla_X e_\beta
\end{aligned} \tag{5.9}$$

Where $X\left(Y^\beta\right)$ is given by,

$$X\left(Y^\beta\right) = X^\alpha e_\alpha\left(Y^\beta\right)$$

Which more specifically can be written using the basis for the tangent space $e_\alpha = \frac{\partial}{\partial x^\alpha}$,

$$X\left(Y^\beta\right) = X^\alpha \frac{\partial}{\partial x^\alpha}\left(Y^\beta\right)$$

Furthermore using (5.8) we can expand out the second term in (5.9).

$$\begin{aligned}
\nabla_X Y &= X\left(Y^\beta\right)e_\beta + Y^\beta \nabla_{X^\alpha e_\alpha} e_\beta && \text{Expand } X \\
&= X\left(Y^\beta\right)e_\beta + Y^\beta X^\alpha \nabla_{e_\alpha} e_\beta \\
&= X\left(Y^\beta\right)e_\beta + Y^\beta X^\alpha \nabla_\alpha e_\beta && \text{By notation} \\
&= X\left(Y^\beta\right)e_\beta + Y^\beta X^\alpha \Gamma^\mu{}_{\alpha\beta} e_\mu && \text{Using the connection } \Gamma \\
&= X^\alpha\left(\partial_\alpha Y^\beta\right)e_\beta + Y^\beta X^\alpha \Gamma^\mu{}_{\alpha\beta} e_\mu && \text{Expand } X \\
\nabla_X Y &= X^\alpha \underbrace{\left(\partial_\alpha Y^\mu + \Gamma^\mu{}_{\alpha\beta} Y^\beta\right)}_{\nabla_\alpha Y^\mu} e_\mu && \text{Relabel indices}
\end{aligned}$$

This means that the components $\nabla_\alpha Y^\mu$ of $\nabla_X Y$ are given by,

$$\nabla_\alpha Y^\mu = \partial_\alpha Y^\mu + \Gamma^\mu{}_{\alpha\beta} Y^\beta \tag{5.10}$$

This is known as the **covariant derivative**. Note that the notation $\frac{\partial y^\mu}{\partial x^\alpha} = \frac{\partial}{\partial x^\alpha} y^\mu = J^\mu{}_\alpha$ is used.

5.5.1 Leibniz Law and Covariant Derivative

The covariant derivative satisfies a product law when dealing with a tensor product. This allows us to extend the notion of a covariant derivative acting on vectors to the covariant derivative acting on a tensor.

$$\nabla_\mu(S \otimes T) = (\nabla_\mu S) \otimes T + S \otimes (\nabla_\mu T)$$

When dealing with dealing with tensors, we will need to know the components of ∇_α acting on a co-vector W_β; $\nabla_\alpha W_\beta$? We should leverage our knowledge of how the covariant derivative acts on a vector Y^β. Let's apply the covariant derivative to the scalar function $Y^\beta W_\beta$ (scalar product). Therefore the covariant derivative should just act like a regular derivative.

$$\nabla_\alpha\left(Y^\beta W_\beta\right) = \partial_\alpha\left(Y^\beta W_\beta\right) = \left(\partial_\alpha Y^\beta\right) W_\beta + Y^\beta(\partial_\alpha W_\beta) \tag{5.11}$$

However, we can also ignore the contraction in β and view the product $Y^\beta W_\beta$ effectively as the tensor product $Y^\gamma W_\beta$ for now. If the covariant derivative is to satisfy a Leibniz law on the tensor product we get,

$$\nabla_\alpha\left(Y^\beta W_\beta\right) = \nabla_\alpha\left(Y^\beta\right) W_\beta + Y^\beta \nabla_\alpha(W_\beta) \tag{5.12}$$

We are aiming to determine the structure of $\nabla_\alpha(W_\beta)$ and we know $\nabla_\alpha\left(Y^\beta\right)$ already as (5.10). Therefore equating (5.11) and (5.12) we get,

$$\left(\partial_\alpha Y^\beta\right) W_\beta + Y^\beta(\partial_\alpha W_\beta) = \nabla_\alpha\left(Y^\beta\right) W_\beta + Y^\beta \nabla_\alpha(W_\beta)$$

Sub in (5.10),

$$\left(\partial_\alpha Y^\beta\right) W_\beta + Y^\beta(\partial_\alpha W_\beta) = \left(\partial_\alpha Y^\beta + \Gamma^\beta{}_{\alpha\mu} Y^\mu\right) W_\beta + Y^\beta \nabla_\alpha(W_\beta)$$

Canceling terms and rearranging,

$$Y^\beta(\partial_\alpha W_\beta) - \Gamma^\beta{}_{\alpha\mu} Y^\mu W_\beta = Y^\beta \nabla_\alpha(W_\beta)$$

Relabel indices for convenience,

$$Y^\beta(\partial_\alpha W_\beta) - \Gamma^\gamma{}_{\alpha\beta} Y^\beta W_\gamma = Y^\beta \nabla_\alpha(W_\beta)$$

$$Y^\beta\{\partial_\alpha W_\beta - \Gamma^\gamma{}_{\alpha\beta} W_\gamma\} = Y^\beta \nabla_\alpha(W_\beta)$$

Since this is true for any Y^β and W_β, we can eliminate Y.

$$\nabla_\alpha(W_\beta) = \partial_\alpha W_\beta - \Gamma^\gamma{}_{\alpha\beta} W_\gamma \tag{5.13}$$

Compare this with (5.10) as there are subtle differences.

Now utilizing the product law we can determine the covariant derivative for any tensor.

$$\begin{aligned}\nabla_\mu T^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} = {}& \partial_\mu T^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} + \cdots \\ & \cdots + \Gamma^{\alpha_1}{}_{\mu\sigma} T^{\sigma\alpha_2\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} + \cdots + \Gamma^{\alpha_p}{}_{\mu\sigma} T^{\sigma\alpha_1\cdots\alpha_{p-1}\sigma}{}_{\beta_1\cdots\beta_q} + \cdots \\ & \cdots - \Gamma^\sigma{}_{\mu\beta_1} T^{\alpha_1\cdots\alpha_p}{}_{\sigma\beta_2\cdots\beta_q} - \cdots - \Gamma^\sigma{}_{\mu\beta_q} T^{\alpha_1\cdots\alpha_1}{}_{\beta_1\cdots\beta_{q-1}\sigma}\end{aligned}$$

It is important to note that $\nabla_\mu T^{\alpha_1\cdots\alpha_p}{}_{\beta_1\ldots\beta_q}$ is still a tensor by *construction*. Since ∂_μ is not a tensor, it *must* be that $\Gamma^\alpha{}_{\mu\sigma}$ is not a tensor. The transformation is written,

$$\Gamma^\alpha{}_{\beta\lambda} \to \Gamma^{\alpha'}{}_{\beta'\lambda'} = \frac{\partial x^\beta}{\partial y^{\beta'}} \frac{\partial x^\gamma}{\partial y^{\gamma'}} \frac{\partial y^{\alpha'}}{\partial x^\alpha} \Gamma^\alpha{}_{\beta\lambda} - \frac{\partial y^{\alpha'}}{\partial x^\gamma} \frac{\partial^2 x^\gamma}{\partial y^{\beta'} \partial y^{\lambda'}}$$

5.6 Geodesics

Definition: The covariant derivative defined in terms of the Christoffel symbol is called the *Levi-Civita* connection.

$$\Gamma^\alpha{}_{\beta\gamma} = \frac{1}{2}g^{\alpha\mu}(g_{\beta\mu,\gamma} + g_{\gamma\mu,\beta} - g_{\gamma\beta,\mu})$$

Where the comma notation indicates a derivative $g_{\alpha\beta,\mu} \equiv \partial_\mu g_{\alpha\beta}$. Note that for the Christoffel symbols Γ, the lower two indices are symmetric $\Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\gamma\beta}$. Under this definition, one can show that the Levi-Civita connection is metric compatible,

$$\nabla_\mu g_{\alpha\beta} = 0$$

Example: For $\mathbb{R}^2$ the Cartesian coordinates the metric $g_{\alpha\beta}$ becomes,

$$g_{\alpha\beta} = \delta_{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This implies that in this example $\Gamma^\alpha{}_{\beta\gamma} = 0$, so clearly $\nabla_\mu g_{\alpha\beta} = 0$ as expected.

Example: Now consider polar coordinates $x^\mu = (r, \theta)$.

$$g_{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix} \to g^{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{bmatrix}$$

Therefore $\Gamma^\alpha{}_{\beta\gamma} \neq 0$, but it is still maintained that $\nabla_\mu g_{\alpha\beta} = 0$. Alternatively, we could have done a change of coordinates to get this result more easily.

Theorem: Let $g_{\alpha\beta}$ be a metric, then there exist a unique torsion free covariant derivative such that $\nabla_\mu g_{\alpha\beta} = 0$. This is the Levi-Civita connection (Proof Wald p.35). A torsion free connection maintains that $\Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\gamma\beta}$.

Note that the Levi-Civita symbol has by product rule,

$$\nabla_\mu\left(g^{\alpha\beta}T_\beta{}^\gamma\right) = g^{\alpha\beta}\nabla_\mu T_\beta{}^\gamma$$

Or that the metric is constant with respect to the Levi-Civita symbol. From now on we replace the derivative ∂_μ with the generalized counterpart ∇_μ. For example, the conservation of the stress energy tensor $\partial_\mu T^{\mu\alpha} = 0$ in Minkowski space and Cartesian coordinates becomes in a general manifold in general coordinates $\nabla_\mu T^{\mu\alpha}$,

$$\partial_\mu T^{\mu\alpha} \to \nabla_\mu T^{\mu\alpha}$$

Applying this to a particle under the influence $\vec{F}_g = m\vec{a}$. Let V^α be the relativistic speed that is a time-like tangent vector, $V^\alpha V_\alpha = -1$.

$$\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau} g_{\alpha\beta} \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = \frac{\mathrm{d}^2 s}{\mathrm{d}\tau^2} = -1$$

No force acting on the particle gives,

$$\frac{\mathrm{d}V^\alpha}{\mathrm{d}\tau} + \Gamma^\alpha{}_{\beta\gamma}V^\beta V^\gamma = 0$$

Re-writing acceleration,

$$V^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau} \to \frac{\mathrm{d}V^\alpha}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\right) = \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau}\frac{\partial}{\partial x^\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}$$

Therefore,

$$V^\beta\partial_\beta V^\alpha + \Gamma^\alpha{}_{\beta\gamma}V^\beta V^\gamma = 0$$
$$V^\beta(\partial_\beta V^\alpha + \Gamma^\alpha{}_{\beta\gamma}V^\gamma) = 0$$
$$V^\beta\nabla_\beta V^\alpha = 0$$

Notationally, this can be compacted as $\nabla_V V^\alpha$. This is known as the **geodesic equation**. Note that ∇_X is just the directional derivative which has been constructed to transform well under a change of coordinates.

Definition: We say that a vector X is parallelly transported along V if $\nabla_V X^\alpha = V^\beta \nabla_\beta X^\alpha = 0$. This **does not** imply that $\nabla_\beta X^\alpha = 0$. Also $\nabla_V g^{\alpha\beta} = V^\gamma \nabla_\gamma g_{\alpha\beta} = 0$ because the inner term $\nabla_\gamma g_{\alpha\beta} = 0$ for fixed γ. This can be extended to any tensor,

$$\nabla_V T^{\alpha_1 \cdots \alpha_p}{}_{\beta_1 \cdots \beta_p} = 0$$

Definition: Consider a curve $\mathcal{C} \subset M$ and V^μ is its tangent vector. If $\nabla_V V^\alpha = 0$ then we call $\mathcal{C}$ a **geodesic**. Geodesics are the generalization of the notion of straight lines.

5.6.1 Examples

Example: Consider $\mathbb{R}^2$ with Cartesian coordinates $x^\mu = (x, y)$ with Minkowski,

$$g_{\alpha\beta} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then a curve $\mathcal{C}$ parameterized by τ is,

$$\mathcal{C} : \begin{bmatrix} x(\tau) \\ y(\tau) \end{bmatrix} \implies V^\alpha = \begin{bmatrix} \dot{x}(\tau) \\ \dot{y}(\tau) \end{bmatrix}$$

Since V^α is time-like, $V^\alpha V_\alpha = -1 = -\dot{x}^2 + \dot{y}^2$ combined with the geodesic equation $\nabla_V V^\alpha = 0$,

$$\frac{\mathrm{d}V^\alpha}{\mathrm{d}\tau} + \underbrace{\Gamma^\alpha{}_{\beta\gamma}}_{=0} V^\beta V^\alpha = 0$$

Thus,

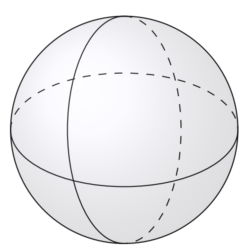
$$\frac{\mathrm{d}V^\alpha}{\mathrm{d}\tau} = \begin{bmatrix} \ddot{x}(\tau) \\ \ddot{y}(\tau) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Which suggests straight lines

$$x^\alpha = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a\tau + b \\ d\tau + e \end{bmatrix}$$

Where a, b, d, e are constants. These are straight lines! Subject to the constraint that $-1 = V^\alpha V_\beta = -a^2 + d^2$.

Example: Geodesic on Sphere,



Consider a sphere with radius r with metric,

$$g_{\alpha\beta} = r^2 \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}$$

With general vector on the sphere x^μ with,

$$x^\mu = \begin{bmatrix} \theta \in [0, \pi] \\ \phi \in [0, 2\pi] \end{bmatrix}$$

Therefore the velocity is V^μ,

$$V^\mu = \begin{bmatrix} \dot{\theta} \\ \dot{\phi} \end{bmatrix}$$

Let our curvilinear parameter of the geodesic be τ and take τ to be the arclength. By doing so, we enforce that $V^\alpha V_\alpha = +1$. Therefore since the arclength $\mathrm{d}\tau^2 = x^\alpha g_{\alpha\beta} x^\beta$,

$$+1 = V^\alpha g_{\alpha\beta} V^\beta = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau} g_{\alpha\beta} \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = \frac{\mathrm{d}\tau^2}{\mathrm{d}\tau^2} = r^2\dot{\theta}^2 + r^2 \sin^2\theta \dot{\phi}^2$$

Therefore the Euclidean metric used gives V^α to be spacelike. This normalization condition $V^\alpha V_\alpha = 1$ is *very* important. Finally, our geodesic must satisfy the geodesic equation.

$$\frac{\mathrm{d}}{\mathrm{d}\tau} V^\alpha + \Gamma^\alpha{}_{\beta\gamma} V^\beta V^\gamma = 0$$

The Christoffel symbols for this metric are given by $(1 = \theta, 2 = \phi)$,

$$\Gamma^2{}_{21} = \Gamma^2{}_{12} = \frac{\cos\theta}{\sin\theta} \qquad \Gamma^1{}_{22} = -\sin\theta\cos\theta$$

Only three terms of the $2^3 = 8$ Christoffel symbols are non-zero. The Geodesic equation becomes,

$$\ddot{\theta} + \Gamma^1{}_{\beta\gamma} V^\beta V^\alpha = 0 \implies \ddot{\theta} - \sin\theta\cos\theta\dot{\phi}\dot{\phi} = 0 \tag{5.14}$$

$$\ddot{\phi} + \Gamma^2{}_{\beta\gamma} V^\beta V^\alpha = 0 \implies \ddot{\phi} + 2\frac{\cos\theta}{\sin\theta}\dot{\theta}\dot{\phi} = 0 \tag{5.15}$$

Noting the very important factor of 2. We can now solve these coupled second order ODEs. First take the case of $\dot{\phi} = 0$ or that ϕ is constant. (5.14) gives $\theta = a\tau + b$. The normalization condition implies that $a = \pm\frac{1}{r}$ in this case of $\dot{\phi} = 0$. This correspond to slices through the sphere at fixed phi. This is a **great circle** or by analogy on the surface of the earth, these are meridians.

Let use examine the other case of θ being constant. In this case, (5.15) indicates that $\ddot{\phi} = 0 \implies \phi = d\tau + e$. Also (5.14) gives,

$$-\sin\theta\cos\theta\left(\dot{\phi}\right)^2 = 0 \implies \dot{\phi}^2 = 0 \text{ or } \sin\theta\cos\theta = 0$$

Note that if $\dot{\phi}^2 = 0$, this forces $d = 0$ which violates the normalization condition. It also corresponds to a fixed point. Alternatively, $\sin\theta\cos\theta = 0$ suggests $\theta = \pi/2 + n\pi$. Note that $\sin\theta \neq 0, \pi$ since this makes (5.15) singular. This solution corresponds to the equator of the sphere. The set of all solutions are all of the great circles on the sphere.

5.6.2 Geodesics & Path Length

When dealing with the Euclidean metric, the geodesics are the straight lines or equivalently the **shortest path**.

$$\text{geodesic} \implies \text{shortest path}$$

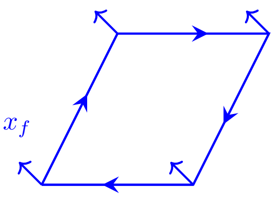
However when dealing with the Lorentzian metric, time like geodesics are the **longest path**.

$$\text{geodesic} \implies \text{longest path}$$

This is to be expected because light-like vectors have a minimum arclength of 0. Therefore in order to make the action of the path have zero variation $\delta S = 0$, the extremized path is one with maximum action S.

5.7 Curvature

Curvature can be examined as a rotation of vectors that are parallel transported around a loop.



When moving around the loop the vector x^μ is not modified.

$$x_i^\mu = x_f^\mu \implies \text{no curvature}$$

Where as for the case of a loop on a sphere the vectors do not match,

$$x_i^\mu \neq x_f^\mu \implies \text{curvature}$$

Definition: Using ∇_μ as the Levi-Civita connection, the *Riemann Curvature Tensor* is defined by,

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) V^\alpha = R_{\mu\nu}{}^\alpha{}_\beta V^\beta$$

This tensor $R_{\mu\nu}{}^\alpha{}_\beta$ encodes the curvature. Expressed in terms of the Christoffel symbols,

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho{}_{\nu\sigma} - \partial_\nu \Gamma^\rho{}_{\mu\sigma} + \Gamma^\rho{}_{\mu\lambda}\Gamma^\lambda{}_{\mu\sigma} - \Gamma^\rho{}_{\nu\lambda}\Gamma^\lambda{}_{\mu\sigma}$$

Notice that when dealing with $4d$ spacetime, the Riemann tensor has $4^4 = 256$ components. It can be viewed as a matrix of matrices. This is a lot of components. Luckily, the symmetries of Γ reduce the number of unique terms. The symmetries are as follows.

$$R_{\rho\sigma\mu\nu} = -R_{\sigma\rho\mu\nu} = -R_{\rho\sigma\nu\mu} \quad \text{Antisymmetric}$$

$$R_{\rho\sigma\mu\nu} = R_{\mu\nu\rho\sigma} \quad \text{Symmetric in Pairs}$$

Also cyclic permutations of the last three indices summed together are zero. This is the *1st Bianchi identity*

$$R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0$$

The *2nd Bianchi identity* deals with cyclic permutations in the first three indices (including ∇_μ),

$$\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\rho R_{\sigma\lambda\mu\nu} + \nabla_\sigma R_{\lambda\rho\mu\nu} = 0 \tag{5.16}$$

The combination of all of these symmetries reduces the number of components from 256 to only 20 independent components.

6 Einstein Field Equations

In the previous section we defined the notion of curvature, $R_{\alpha\beta\nu\mu}$ which is a function of $(\partial\Gamma, \Gamma)$. Also, the Christoffel symbols are a function of $(\partial g, g)$. We are looking for a generalization of $\vec{F}_g = m_I \vec{a}$. Our analysis of $\Delta\phi = 4\pi G\rho$ lead to the suggestion that $g_{\mu\nu}(\phi)$; the metric was dependent on ϕ.

6.1 Motivation

Let us take Poisson's equation to get,

$$\Delta\phi = 4\pi G\rho$$

The constants $4\pi G$ will be generalized to constants k, and ρ generalizes to the stress energy tensor $T_{\alpha\beta}$. How does $\Delta\phi$ modify?

$$f(R_{\alpha\beta\mu\nu}) \stackrel{?}{=} kT_{\alpha\beta}$$

The RHS of this equation is a rank $(0, 2)$ tensor. Therefore we will need to contract some of the indices of $R_{\alpha\beta\mu\nu}$. Note that we can not just multiply it by the metric because as a consequence of symmetry,

$$g^{\alpha\beta}R_{\alpha\beta\mu\nu} = 0 \qquad g^{\mu\nu}R_{\alpha\beta\mu\nu} = 0$$

Therefore we will contract against the first indices of each pair,

$$g^{\alpha\mu}R_{\alpha\beta\mu\nu} \equiv R_{\beta\nu}$$

This is known as the **Ricci tensor**. Note that by extension of properties of the curvature,

$$R_{\beta\nu} = R_{\nu\beta}$$

Thus Einstein's first proposal is,

$$R_{\beta\nu} = kT_{\beta\nu}$$

This is well motivated as both the right and left hand sides are symmetric. However since the stress energy tensor is conserved,

$$\nabla^{\nu}T_{\beta\nu} = 0$$

But is can be shown using the second Bianchi identity (5.16),

$$\nabla^{\nu}R_{\nu\mu} = \frac{1}{2}\nabla_{\mu}R$$

Where $R = R_{\alpha\beta}g^{\alpha\beta}$ is known as the **Ricci scalar**. Therefore this first proposal is not consistent.

How can we modify this to maintain consistency? Define a new quantity,

$$G_{\beta\nu} = R_{\beta\nu} - \frac{1}{2}Rg_{\beta\nu}$$

Such that it is divergenceless.

$$\begin{aligned}
\nabla^{\nu}G_{\nu\beta} &= \nabla^{\nu}R_{\nu\beta} - \frac{1}{2}\nabla^{\beta}(Rg_{\beta\nu}) \quad \text{Definition of } G_{\nu\beta} \\
&= \nabla^{\nu}R_{\nu\beta} - \frac{1}{2}\nabla^{\beta}(R)g_{\beta\nu} - \frac{1}{2}R\nabla^{\beta}(g_{\beta\nu}) \quad \text{Product rule} \\
&= \nabla^{\nu}R_{\nu\beta} - \frac{1}{2}g_{\beta\nu}\nabla^{\beta}R \quad \text{Metric compatible} \\
&= \frac{1}{2}\nabla_{\nu}R - \frac{1}{2}g_{\beta\nu}\nabla^{\beta}R \\
&= \frac{1}{2}\nabla_{\nu}R - \frac{1}{2}\nabla_{\nu}R \quad \text{Lowering index} \\
&= 0
\end{aligned}$$

However, is this G unique? Obviously we can add any constants that are compatible with the metric. Therefore introduce the **cosmological constant** Λ.

$$G_{\beta\nu} = R_{\beta\nu} - \frac{1}{2}Rg_{\beta\nu} + \Lambda g_{\beta\nu}$$

Therefore we have $G_{\beta\nu} = G_{\nu\beta}$ and that $\nabla^{\nu}G_{\beta\nu} = 0$. Therefore,

$$G_{\beta\nu} = kT_{\beta\nu}$$

Where k is yet to be determined.

To determine what k is, we should be able to recover Newton's law of gravity in a non-relativistic limit. Consider a weak field stationary metric.

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} = \eta_{\alpha\beta} - 2\phi\delta_{\alpha\beta}$$

Where $\frac{\partial}{\partial x^\beta} h_{\alpha\beta} = 0$. Now we can calculate $T_{\alpha\beta}$. Consider a perfect fluid as a dust with pressure $P = 0$ and mass density $\rho = 0$.

$$T^{\alpha\beta} = \rho V^\alpha V^\beta$$

For $h_{\alpha\beta} \ll 1$ or in a weak gravitational field, a little bit of work yields,

$$R_{\alpha\beta} = -\frac{1}{2}\Box h_{\alpha\beta}$$

In the limit that $c \gg 1$ we obtain that the D'Alembertian becomes the Laplacian.

$$R_{\alpha\beta} = \Delta\phi\delta_{\alpha\beta}$$

Therefore we obtain (assuming $\Lambda = 0$ since we are dealing with Newtonian physics),

$$G_{\alpha\beta} = kT_{\alpha\beta}$$

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = kT_{\alpha\beta} \tag{6.1}$$

Now we take the trace of this equation.

$$R - \frac{1}{2}Rg_{\alpha\beta}g^{\alpha\beta} = kT_{\alpha\beta}g^{\alpha\beta}$$

Notice that $g_{\alpha\beta}g^{\alpha\beta} = 4$, since we are in a 4 dimensional spacetime.

$$g^{\alpha\beta}g_{\beta\gamma} = \delta^\alpha{}_\gamma$$

If we explicitly sum up this product,

$$g^{\alpha\beta}g_{\beta\alpha} = \delta^\alpha{}_\alpha$$

We obtain the *trace* of the identity matrix which yields 4.

$$-R = kT_{\alpha\beta}g^{\alpha\beta} = kT$$

Therefore we can write (6.1) as,

$$R_{\alpha\beta} = kT_{\alpha\beta} + \frac{1}{2}Rg_{\alpha\beta}$$

$$R_{\alpha\beta} = kT_{\alpha\beta} - \frac{1}{2}kTg_{\alpha\beta}$$

Switching to the rest frame of the fluid, we obtain $T_{\alpha\beta} = \rho V^\alpha V^\beta$ as,

$$T_{\alpha\beta} = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The only remaining component is $T_{00} = \rho$.

$$R_{00} = kT_{00} - \frac{1}{2}kT_{00}g^{00}g_{00}$$

Notice that $g^{00}g_{00}$ is given by,

$$g_{00} = -1 - 2\phi \qquad g^{00} = -1 + 2\phi$$

$$g^{00}g_{00} = 1 + O(\phi^2) \approx 1$$

Therefore,

$$R_{00} = \frac{1}{2}kT_{00} = \frac{1}{2}k\rho$$

$$+\Delta\phi = \frac{1}{2}k\rho$$

Which restricts k to be $k = 8\pi G$. Finally we obtain,

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} + \Lambda g_{\alpha\beta} = 8\pi GT_{\alpha\beta}$$

This is **Einstein's field equation**. This acts as our generalization of $\Delta\phi = 4\pi G\rho$.

6.2 Recap of General Relativity

As a recap on everything we have learned so far,

1. gravitational degrees of freedom $\phi \to g_{\alpha\beta}$ metric on the manifold which acts as spacetime
2. given that we known the metric, we can determine completely the Levi-Civita connection
   - $\nabla_\mu V^\alpha = \partial_\mu V^\alpha + \Gamma^\alpha{}_{\beta\gamma}V^\gamma$
   - $\Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\gamma\beta}$
   - $\nabla_\mu g_{\alpha\beta} = 0$
   - from now on we only consider the Levi-Civita connection
3. geodesics are the paths of particles in spacetime subject to **no other forces**
   - $\nabla_V V^\alpha = 0$ where V^α is the tangent vector to the geodesic
   - τ: arclength with $V^\alpha V_\alpha = -1, 0, 1$
   - $\frac{\mathrm{d}V^\alpha}{\mathrm{d}\tau} + \Gamma^\alpha{}_{\beta\gamma}V^\beta V^\gamma = 0$ System of differential equations
   - Notice that there is no mass present in the geodesic equation
   - when forces are applied to a particles with mass m, $m\left\{\frac{\mathrm{d}V^\alpha}{\mathrm{d}\tau} + \Gamma^\alpha{}_{\beta\gamma}V^\beta V^\gamma\right\} = F^\alpha$
   - when there is a force it is **no longer a geodesic**
4. Stress energy tensor $T^{\alpha\beta}$
   - represents a mass/energy density and is derived from a flux of particles
   - symmetric $T^{\alpha\beta} = T^{\beta\alpha}$
   - conservation $\nabla_\alpha T^{\alpha\beta} = 0$ (equation of motion for fluid)
   - $\nabla_\alpha T^{\alpha\beta} = 0$ gives the continuity equation and Navier-Stokes equation in non-relativistic limit
5. Covariant derivative can be used to define curvature
   - Riemann tensor $[\nabla_\mu, \nabla_\nu]V^\alpha = R_{\mu\nu}{}^\alpha{}_\beta V^\beta$ ($R_{\mu\nu}{}^\alpha{}_\beta = 0$ implies zero curvature)
   - Ricci tensor $R_{\alpha\beta} = R_{\mu\alpha\gamma\beta}g^{\mu\gamma}$ (not that $R_{\alpha\beta} = 0$ does not imply that there is zero curvature)
   - Ricci scalar $R = R_{\alpha\beta}g^{\alpha\beta}$
6. Einstein equations
   - $G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = 8\pi GT_{\alpha\beta}$ (if there is a source of matter and energy)
   - $G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = 0$ (in vacuum)
   - $G_{\alpha\beta} = G_{\beta\alpha}$

- $\nabla^\alpha G_{\alpha\beta} = 0$

7. Einstein equivalence principle

   - There exists a locally inertial coordinate system ξ^μ *at the point* p such that $g_{\bar{\alpha}\bar{\beta}}(p) = \eta_{\bar{\alpha}\bar{\beta}}(p)$
   - Physics is like in Minkowski space time at p

The equivalence principle holds together all of General relativity,

$$M, g_{\alpha\beta} \to \nabla_\mu, \text{geodesic eqn} \to \text{SET } T_{\alpha\beta} \to \text{curvature } R_{\mu\nu\alpha\beta} \to \text{Einstein equation}$$

6.3 Alternative Approaches to Gravity

Visualized in this way, general relativity becomes quite simple. How can we make it more complicated? Suppose we tried the followed,

$$\phi \to g_{\alpha\beta} + V_\alpha + M_{\alpha\beta\gamma\delta}$$

By combing tensor theories $g_{\alpha\beta}$ and vector theories V_α. Maybe these theories of gravity can explain dark matter. This approach adds numerous degrees of freedom. Alternatively, one could include torsion into the connections and move away from the Levi-Civita connection. Moreover, we could add more dimensions.

In each of these cases, we must compare out observations/experiments with the predictions of the theory. These observations can be made on the three scales,

Scale	**GR**
Solar system	works very well
Galactic	issues with dark matter
Cosmological	$\Lambda = ?$

6.4 Solving Einstein's Equation

Einstein's field equations are difficult to solve. They are a system of coupled non-linear PDEs.

$$G_{\alpha\beta} = kT_{\alpha\beta} \stackrel{?}{=} 0$$

It is still an active area of research to solve these equations. For a point-like particle, the particles follows a geodesic,

$$\nabla_V V^\alpha = 0$$

However if the particle has mass, it *should* affect the spacetime it occupies and thus the metric. This affect of the mass affecting directly the space it moves in is also an area of active research. In principle $T_{\alpha\beta}$ depends on the trajectory of the particles V,

$$G_{\alpha\beta} \stackrel{?}{=} kT_{\alpha\beta}(V)$$

6.4.1 Assumptions to Simplify Einstein's Equations

As a physicist, a non-linear equation is difficult to solve. We will linearize the metric using perturbations such as $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$ where $h_{\alpha\beta}$ characterizes the dynamics and propagation of gravity (gravitational waves).

Additionally, symmetries are going to greatly simplify the equations. These symmetries will be exposed using killing vectors.

6.5 Schwarzschild Metric

For example, as in Newtonian gravity we can consider a system that has spherical symmetry and time translational invariance. The system will have mass M. First we will consider the approach from a Newtonian perspective and then compare and contrast with the metric approach. Inside the system,

$$\Delta\phi = 4\pi G\rho$$

Whereas outside the system, since there is no mass,

$$\Delta\phi = 0$$

Use spherical coordinates,

$$\Delta\phi = \frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}}{\mathrm{d}r}\phi(r)\right)$$

Outside the system, ϕ must be like in general with $A, B \in \mathbb{R}$ constants with A having dimension length,

$$\phi(r) = -\frac{A}{r} + B$$

The condition at $r \to \infty$ demanding that $\lim_{r\to\infty}\phi(r) = 0$ enforces $B = 0$. We also want to experience an attractive force where $A > 0$. Evidently, we should expect the potential ϕ to depend on M . In order to give A dimensions of length, we will need to multiply by G.

$$\phi = -\frac{GM}{r}$$

Notice that we didn't explicitly make use of the time translational invariance here.

Now let us attempt this problem from the general relativity approach. We generalize $\phi \to g_{\alpha\beta}$ and $g_{\alpha\beta}$ should be spherically symmetric. The time translational invariance gives,

$$g_{\alpha\beta}(r,\theta,\phi,t) \to g_{\alpha\beta}(r,\theta,\phi)$$

So what is the most general metric we can write with these restrictions,

$$g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = -e^{2\alpha(r)}\mathrm{d}t^2 + e^{2\beta(r)}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2$$

Where $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2$ is the angular element and $\alpha(r), \beta(r)$ are functions of r and r alone. Note the exponentials are a manifestation of the fact that the metric should be *Lorentzian* (signature $(-,+,+,+)$) which enforces a strictly negative coefficient in front of $\mathrm{d}t^2$. Comparing to the Minkowski metric,

$$\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2$$

Note that both of these metrics are diagonal. What about terms $\mathrm{d}t\mathrm{d}x^\mu$? These can not be present due to the time invariance.

Outside the system with no sources, Einstein's equation becomes $G_{\mu\nu} = 0$. Therefore,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0$$

From here, one can calculate $R_{\mu\nu}$ and R from the definition of the metric. Then solving the resultant differential equations to get $\alpha(r)$ and $\beta(r)$. For example,

$$R_{tt} = e^{2(\alpha-\beta)}\left(\partial_r^2\alpha + (\partial_r\alpha)^2 - \partial_r\alpha\partial_r\beta + \frac{2}{r}\partial_r\alpha\right)$$

The rest of the terms are similarly messy. Solving these equations gives,

$$\alpha = -\beta \qquad \partial_r\left(re^{2\alpha}\right) = 1$$

Therefore with R_s a constant to be determined by boundary conditions,

$$e^{2\alpha} = 1 - \frac{R_s}{r}$$

The metric becomes,

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -\left(1 - \frac{R_s}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{R_s}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2$$

This metric is known as the **Schwarzschild metric**.

Binkhoff Theorem:
The Schwarzschild metric (other that Minkowski) is the *unique* vacuum solution with *spherical symmetry*.

Now let use examine the case of $\frac{R_s}{r} \ll 1$ the distance from the source is very large.

$$\left(1 - \frac{R_s}{r}\right)^{-1} \approx \left(1 + \frac{R_s}{r}\right)$$

Then the far field metric is,

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \approx -\left(1 - \frac{R_s}{r}\right)\mathrm{d}t^2 + \left(1 + \frac{R_s}{r}\right)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2$$

This resembles the *weak field metric* discussed earlier. If we set $\phi(r) = -\frac{1}{2}\frac{R_s}{r}$,

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \approx -(1 + 2\phi)\mathrm{d}t^2 + (1 - 2\phi)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu + h_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$$

With $h_{\alpha\beta} = -\phi(r)\delta_{\alpha\beta}$. From this we can identify the value of R_s,

$$\phi(r) = -\frac{1}{2}\frac{R_s}{r} = -\frac{GM}{r} \implies R_s = 2GM$$

For the sun, $R_s = 3\,\mathrm{km}$ and for the earth, $R_s - 8.7\,\mathrm{mm}$. In both cases, these are much less than the radius of the object $R_s \ll R$. As such, the earth and the sun are not black holes.

If we are very far away, $r \gg 1$, the Schwarzschild metric in the limit becomes the weak-field metric and even further becomes the Minkowski metric. As such we say that the Schwarzschild metric is **asymptotically flat**. What about the case of $r = R_s$? Here there is a *coordinate singularity*, which happens to be an event horizon. Note that this is not a true singularity in the sense that is coordinate dependent. The true singularity lives at $r = 0$. We will return to this later.

Example: Consider a system that generates a Schwarzschild metric. We can study this system's physics. First notice that the Schwarzschild metric is time independent. Secondly, the Schwarzschild metric is a very particular solution to Einstein's equations. Near large spherical objects like the sun, the Schwarzschild metric is a good approximation (ignoring the effects of other bodies in the universe, as well as feedback loops between two bodies), but in general the metric of the universe is much more complicated.

- What is the redshift of this object/system?
- What are the orbits of particles (modeling planets)?
- What is the trajectory of light in this metric? What is the bending of light?
- Geodesic precession?

Each of these bits of physics deals with geodesics. The last thing to consider is the manifestation of **black-holes**.

6.6 Red Shift

In order to compute red shift it is convenient to utilize **killing vectors**; see Appendix C. Since the Schwarzschild metric does not depend on time,

$$\partial_t g_{\alpha\beta} = 0$$

As such, we have a time killing vector ξ^μ,

$$\xi_\mu = (1, 0, 0, 0)$$

In the rest frame of the observer we have,

$$u^\alpha = \left(u^0, 0, 0, 0\right)$$

Since in the rest frame of the observer, the observer is a geodesic, we have the normalization condition, $u^\alpha g_{\alpha\beta} u^\beta = -1$.

$$u^0 = \pm\left(1 - \frac{R_s}{r}\right)^{-1/2}$$

We will take the positive solution since the particle is moving forward in time. Notice now that V^α is collinear with ξ^α,

$$u^\alpha = \left(1 - \frac{R_s}{r}\right)^{-1/2} \xi^\alpha$$

The energy of a photon with momentum vector p^α relative to observer with tangent u^α is given by the scalar product within the metric,

$$E = u^\alpha p_\alpha$$

Naturally, in the frame of the observer, u^α acts as a basis for the time-like vectors. This is just the projection of the photon momentum on the observer.

So what is the red shift of the photon as it moves from radius r_1 to r_2? Since ξ^α is a killing vector, $\xi^\alpha p_\alpha$ is conserved on the photon's world line.

$$\nabla_V(\xi^\alpha p_\alpha) = \nabla_V(\xi^\alpha) p_\alpha + \xi^\alpha \underbrace{\nabla_V(p_\alpha)}_{=0}$$

Where $\nabla_V(p_\alpha)$ is zero because it is a geodesic. The remaining term is,

$$\nabla_V(\xi^\alpha p_\alpha) = V^\gamma \nabla_\gamma(\xi^\alpha) p_\alpha = V^\gamma \nabla_\gamma(\xi_\alpha) g^{\alpha\beta} p_\beta = V^\gamma V^\beta \nabla_\gamma(\xi_\alpha) =?$$

But by the killing equation,

$$\nabla_\gamma \xi_\beta + \nabla_\beta \xi_\gamma = 0$$

Which indicates that $\nabla_\gamma \xi_\beta$ is an antisymmetric tensor. Noting that $V^\alpha V^\beta$ is clearly a symmetric tensor. By contracting a symmetric tensor with an antisymmetric tensor,

$$\nabla_V(\xi^\alpha p_\alpha) = 0$$

If $\xi^\alpha p_\alpha$ is conserved along the geodesic,

$$E(r) = u^\alpha p_\alpha = (1 + 2\phi(r))^{-1/2} \xi^\alpha p_\alpha$$

At $r = r_1$ we have,

$$E(r_1) = (1 + 2\phi(r_1))^{-1/2} \xi^\alpha p_\alpha$$

Identically for r_2,

$$E(r_2) = (1 + 2\phi(r_2))^{-1/2} \xi^\alpha p_\alpha$$

Since $\xi^\alpha p_\alpha$ is the same quantity at both r_1 and r_2. Taking a ratio of the energies,

$$\frac{E_1}{E_2} = \frac{(1+2\phi(r_1))^{-1/2}}{(1+2\phi(r_2))^{-1/2}}$$

Performing a Taylor series in the weak field limit $\phi \ll 1$ OR equivalently $R_s \gg 1$,

$$\frac{E_1}{E_2} \approx 1 + \frac{R_s}{2}\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

Recall that $E = hc/\lambda$. Therefore the wavelength ratio is given by,

$$\frac{\lambda_2}{\lambda_1} \approx 1 + \frac{R_s}{2}\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

Since $r_2 > r_1$, this implies that $\lambda_2 > \lambda_1$. This effect is red shift! The photon loses energy as it climbs out of the gravitational field.

6.7 Orbits & Precession

Now we will focus on the trajectories of particles in the Schwarzschild metric. What we hope to recover is **perihelion precession** of orbits. The *perihelion* is in closest point to the sun in the orbit, the *aphelion* is the corresponding farthest point. Perihelion precession is the rotation of this point around the sun $\Delta\phi$. Newtonian gravity can not explain this effect and no other planets have been observed to explain this effect for Mercury's precession.

The orbits of planets will be taken as orbits of particles,

$$\text{planet} = \text{particle}$$

The orbit of a particle will be a time-like geodesic with helical spiral traces out in a spacetime diagram. The normalization for a time-like curve is,

$$g_{\alpha\beta} V^\alpha V^\beta = -1$$

Where $V^\alpha = \left(\dot{t}, \dot{r}, \dot{\theta}, \dot{\phi}\right)$. The normalization condition is explicitly,

$$-\left(1 - \frac{R_s}{r}\right)\dot{t}^2 + \left(1 - \frac{R_s}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2 = -1 \tag{6.2}$$

The geodesic equations on the other hand are more complicated,

$$\begin{aligned}
\dot{t} + \frac{R_s}{r(r-R_s)}\dot{r}\dot{t} &= 0 \\
\ddot{r} + \cdots &= 0 \quad \text{Won't be used} \\
\frac{\mathrm{d}}{\mathrm{d}\tau}\left(r^2\dot{\theta}\right) - r^2\sin\theta\cos\theta\left(\dot{\phi}\right)^2 &= 0 \\
\ddot{\phi} + \frac{2}{r}\dot{\phi}\dot{r} + 2\frac{\sin\theta}{\cos\theta}\dot{\theta}\dot{\phi} &= 0
\end{aligned} \tag{6.3}$$

Where the factor of two in the last equation are due to the symmetric nature of the Christoffel symbols $\Gamma^\phi{}_{\theta\phi}$. We will make use of conserved quantities. The metric is time invariant meaning we have a time killing vector $\xi^\alpha = (1, 0, 0, 0)$. Also, the metric is spherically symmetric. From this we should expect 3 killing vectors. Since the metric is independent of ϕ one of the killing vectors is ∂_ϕ which in components is $\xi^\alpha = (0, 0, 0, 1)$. This ϕ killing vector will be denoted R_ϕ. The other two are not as simple. Therefore along the geodesic we have conserved quantities,

$$E = \xi^\mu V_\mu = \xi^\mu g_{\mu\alpha} V^\alpha$$

Where V^μ is the tangent vector of the geodesic. Up to the mass of the particle, this is just the energy as before. Using the metric explicitly,

$$E = \xi^0 g_{0\alpha} V^\alpha = g_0 0 V^0 = -\left(1 - \frac{R_s}{r}\right)\dot{t} \tag{6.4}$$

The other conserved quantity will be one component of the angular momentum.

$$L = R^\mu_\phi V_\mu = R^\mu_\phi g_{\mu\alpha} V^\alpha = g_{\phi\phi} V^\phi = r^2 \sin\theta \dot{\phi}$$

In order to expose an obvious solution, $\theta(\tau) = \frac{\pi}{2}, \forall \tau$. This trivially solves (6.3). This implies that $L = r^2\dot{\phi}$. Focusing at the normalization condition (6.2). Multiply (6.2) by $\left(1 - \frac{R_s}{r}\right)$. The first term becomes E^2 by (6.4), and the other terms become L^2/r^2,

$$-E^2 + \dot{r}^2 + \frac{L^2}{r^2}\left(1 - \frac{R^s}{r}\right) = -\left(1 - \frac{R_s}{r}\right)$$

Notice that this is only a differential equation in r. Rearranging yields,

$$\frac{1}{2}\dot{r}^2 + V(r) = \mathcal{E}$$

Where $V(r)$ is given by,

$$V(r) = \frac{1}{2} - \frac{1}{2}\frac{R_s}{r} + \frac{L^2}{2r^2} - \frac{1}{2}\frac{R_s L^2}{r^3}$$

and $\mathcal{E} = \frac{1}{2}E^2$. We will now multiply this equation by $\dot{\phi}^{-2}$. After come algebra and defining $x = \frac{2L^2}{R_s r}$ and differentiate with respect to ϕ (since r will be a function of ϕ) we get a second order differential equation,

$$\frac{\mathrm{d}^2 x}{\mathrm{d}\phi^2} - 1 + x = \frac{3}{4}\frac{R_s^2}{L^2}x^2 \tag{6.5}$$

In the Newtonian cases, this corresponding equation of motion is simpler,

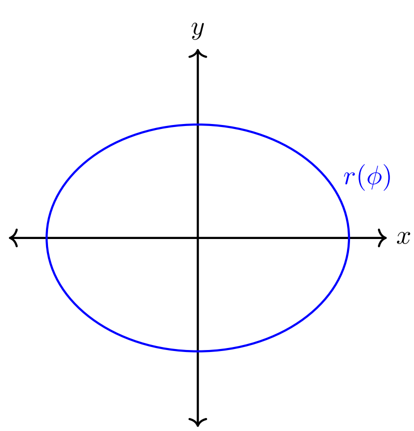
$$\frac{\mathrm{d}^2 x}{\mathrm{d}\phi^2} - 1 + x = 0$$

This extra term $\frac{3}{4}\frac{R_s^2}{L^2}x^2$ is responsible for precession of the orbits. In order to solve (6.5), we will first consider the homogeneous equation and then solve (6.5) by perturbations. The solution to the homogeneous case is given by,

$$x = 1 + e\cos\phi \tag{6.6}$$

Where e is an undetermined constant known as the eccentricity of the ellipse. Using the definition of r and x gives,

$$\frac{2L^2}{R_s r} = 1 + e\cos\phi \implies r = \frac{2L^2}{R_s(1 + e\cos\phi)}$$



To solve (6.5), we will use x_0 as the solution to the homogeneous differential equation (6.6) and perturb the solution by x_1.

$$x = x_0 + x_1$$

Subbing into (6.5) gives,

$$\frac{\mathrm{d}^2 x_1}{\mathrm{d}\phi^2} + x_1 = \frac{3}{4}\frac{R_s^2}{L^2}(x_0 + x_1)^2$$

Treating $x_0 + x_1 \approx x_0$ since x_1 is much less than x_0,

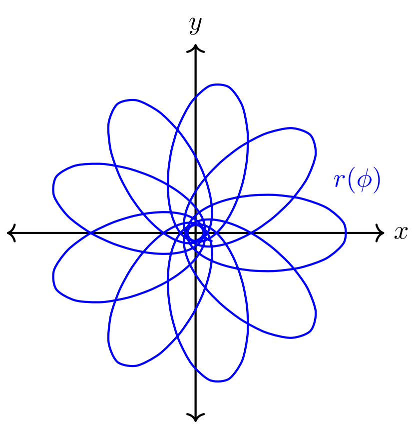
$$\frac{\mathrm{d}^2 x_1}{\mathrm{d}\phi^2} + x_1 \approx \frac{3}{4}\frac{R_s^2}{L^2}x_0^2$$

This gives solution,

$$x \approx 1 + e\cos((1-\alpha)\phi)$$

With $\alpha = \frac{3}{4}\frac{R_s^2}{L^2}$. Therefore the moment ϕ goes completely around one revolution $\phi = 2\pi$, the actual orbit has gone around $2\pi(1-\alpha)$ times. This means that the orbit has not done a complete revolution. This manifests as a precession of the orbit. The perihelion advances by $\Delta\phi = 2\pi\alpha$.

$$\Delta\phi = \frac{3}{4}\pi\frac{R_s^2}{L^2} \approx \frac{3\pi R_s}{(1-e^2)a}$$



For Mercury, this effect is measurable but still small.

$$\Delta\phi = 43'' \text{ per century}$$

6.8 Deflection of Light

As we have seen the geodesic equation gives us a solution for $\theta(\tau) = \pi/2$. When considering a *light-like* vector now, the normalization for this geodesic was given to be,

$$-\left(1 - \frac{R_s}{r}\right)\dot{t}^2 + \left(1 - \frac{R_s}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2 = 0$$

Notice the contrast to (6.2). Taking again $\theta(\tau) = \pi/2 \forall\tau$, we have conserved quantities,

$$E = -\left(1 - \frac{R_s}{r}\right)\dot{t} \qquad L = +r^2\sin^2\theta\dot{\phi} = r^2\dot{\phi}$$

Which simplifies the normalization condition,

$$-\left(1 - \frac{R_s}{r}\right)^{-1}E^2 + \left(1 - \frac{R_s}{r}\right)^{-1}\dot{r}^2 + \frac{L^2}{r^2} = 0$$

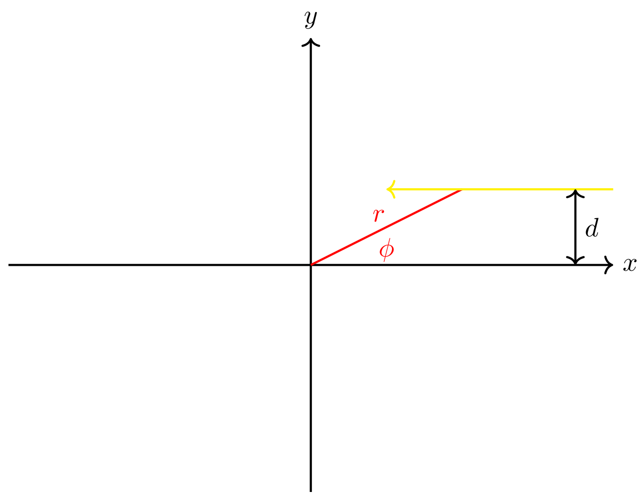
Multiplying by $-\left(1-\frac{R_s}{r}\right)/L^2$

$$\frac{E^2}{L^2}-\frac{\dot{r}^2}{L^2}-\left(1-\frac{R_s}{r}\right)\frac{1}{r^2}=0$$

With $b=\frac{L}{E}$ and $W(r)=\left(1-\frac{R_s}{r}\right)\frac{1}{r^2}$,

$$\frac{1}{b^2}=\frac{\dot{r}^2}{L^2}+W(r)=0$$

Now that this differential equation for r has been simplified, we can solve for r and get 3 distinct types of trajectories. The first of the these trajectories is a **circular orbit** with radius $r=3GM$. The second trajectory is the light being captured by the gravitational well and is an **inward spiral** that terminates. The final trajectory is the light being **deflected** when moving past the body. We will focus now on the deflection case.



When light is very far away, $d/r=\sin\phi\approx\phi$. If $\frac{\mathrm{d}\phi}{\mathrm{d}r}=-d/r^2$ what is $\frac{\mathrm{d}\phi}{\mathrm{d}t}$? If their is no deflection, the change in ϕ, $\Delta\phi$ should be π. For deflection,

$$\Delta\phi=\pi+\delta\phi$$

For the Geodesic,

$$\frac{1}{b^2}+\frac{1}{L^2}\dot{r}^2+W_{\text{eff}}(r) \tag{6.7}$$

With $b=L/E$ a conserved quantity. Therefore,

$$b^2=\frac{L^2}{E^2}=\frac{r^4\dot{\phi}^2}{\left(1-\frac{R_s}{r}\right)^2\dot{t}^2}=\frac{r^4}{\left(1-\frac{R_s}{r}\right)^2}\left(\frac{\mathrm{d}\phi}{\mathrm{d}t}\right)^2$$

Next we can take $r\gg R_s$ to give the approximation,

$$b^2\approx r^4\left(\frac{\mathrm{d}\phi}{\mathrm{d}t}\right)^2$$

We also know that for r very large, the Schwarzschild metric becomes approximately the Minkowski metric. If we have $\sin\phi=d/r$, a large r indicates that ϕ is small with,

$$\phi=\frac{d}{r}$$

What then is $\frac{\mathrm{d}\phi}{\mathrm{d}t}$? By chain rule we get,

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}=\frac{\mathrm{d}\phi}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}t}=-\frac{d}{r^2}\frac{\mathrm{d}r}{\mathrm{d}t} \tag{6.8}$$

What then is $\frac{\mathrm{d}r}{\mathrm{d}t}$? By the normalization condition for a light curve is $V^\alpha V_\alpha = 0$. At large r again,

$$0 = -\dot{t}^2 + \dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2$$

Fixing $\theta\tau = \pi/2, \forall\tau$ again,

$$0 = -\dot{t}^2 + \dot{r}^2 + r^2\dot{\phi}^2$$

Which gives,

$$1 = \left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + r^2\left(\frac{\mathrm{d}\phi}{\mathrm{d}t}\right)^2$$

By substituting (6.8), we get,

$$1 = \left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2\left(1 + \frac{d^2}{r^4}r^2\right)$$

For large r again,

$$1 \approx \left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2$$

Since the light trajectory has r decreasing,

$$\frac{\mathrm{d}r}{\mathrm{d}t} = -1$$

As such, we arrive at the result using (6.8),

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = +\frac{d}{r^2}$$

Which gives,

$$b^2 = r^4\left(\frac{\mathrm{d}\phi}{\mathrm{d}t}\right)^2 = r^4\left(\frac{d}{r^2}\right)^2 = d^2 \implies b = d$$

The conserved quantity b indicates that d is also constant and equal to b. By (6.7),

$$\frac{1}{d^2} = \frac{\dot{r}^2}{L^2} + W_{\text{eff}}(r)$$

With $L^2 = r^4\dot{\phi}^2$,

$$\frac{1}{d^2} = \frac{1}{r^4}\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 + W_{\text{eff}}(r)$$

This is just a differential equation for r and ϕ. Inverting gives,

$$\left(\frac{\mathrm{d}\phi}{\mathrm{d}r}\right)^2 = \frac{1}{r^4}\left(\frac{1}{d^2} - W_{\text{eff}}(r)\right)^{-1}$$

When taking this equation's square root, should be retain the positive or the negative result. When the light ray is initially approaching (say in region 1), an increasing ϕ means a *decreasing* r. Albeit when the light ray is leaving the origin (say in region 2), an increasing ϕ means an *increasing* r. Therefore we have both regions to consider.

$$\frac{\mathrm{d}\phi}{\mathrm{d}r} = \pm\frac{1}{r^2}\left(\frac{1}{d^2} - W_{\text{eff}}(r)\right)^{-1/2}$$

Taking r_1 to the be the point of transition from region 1 to region 2,

$$\Delta\phi = \int_\infty^{r_1} -\frac{1}{r^2}\left(\frac{1}{d^2} - W_{\text{eff}}(r)\right)^{-1/2}\mathrm{d}r + \int_{r_1}^\infty \frac{1}{r^2}\left(\frac{1}{d^2} - W_{\text{eff}}(r)\right)^{-1/2}\mathrm{d}r$$

$$\Delta\phi = 2\int_{r_1}^\infty \frac{1}{r^2}\left(\frac{1}{d^2} - W_{\text{eff}}(r)\right)^{-1/2}\mathrm{d}r$$

After integrating,

$$\Delta\phi = \pi + \delta\phi$$

With $\delta\phi = 2\frac{R_s}{d} = \frac{4GM}{d}$. For the sun and rays grazing the edge of the sun,

$$\delta\phi_\odot \approx 1.7''$$

Such an effect would also appear in Newtonian gravity but off by a factor of 2,

$$\delta\phi_{\text{Newtonian}} = \frac{1}{2}\delta\phi_{\text{GR}}$$

In 1919, Eddington measured the location of the stars behind the sun during an solar eclipse compared to where they are normally. This confirmed the theory of GR and made Einstein a **rock-star**.

Due to the bending of light, **Gravitational lensing** allows for the bending of light around large massive bodies. This allows one to see behind the massive body. Alternatively, measuring a deflection $\delta\phi$ and the mass M and size d of large clusters of objects, many times the predictions of GR are **wrong**. There are many proposals to explain this discrepancy. Since this matter has not been found by any other means, one would expect it to only interact with gravity and gravity alone. As such, it is called **dark matter**. In practice, we need to add *extra* mass in order to match the actual deflection.

Another hint that indicates something might be wrong with gravity, rotating galaxies should have their exterior arms rotate slower as they are farther away from the Galactic center. Unfortunately, with increasing r, the rotational velocity does not decrease; instead is remains constant. This suggests there is either Dark Matter halos or gravity is wrong.

6.9 Black Holes

For the sun, the Schwarzschild radius is only 3 km while its actual radius of the sun is $R_{\text{sun}} = 7 \times 10^5$ km. If the radius was shrunk below the Schwarzschild radius, nothing bad would happen *gravitationally.* The Schwarzschild metric is valid *outside* the spherically symmetric object. Recall the troubling term in the Schwarzschild metric,

$$\left(1 - \frac{R_s}{r}\right)^{-1}$$

Which makes the metric diverge at $r = R_s$,

$$\mathrm{d}s^2 = -\left(1 - \frac{R_s}{r}\right)\mathrm{d}t^2 + \underbrace{\left(1 - \frac{R_s}{r}\right)^{-1}}_{\to\infty}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \tag{6.9}$$

Because of this, the Schwarzschild metric only works in th region with $r > R_s$. The Schwarzschild metric is a solution to the vacuum Einstein equations. Remember that,

$$G_{\mu\nu} = 0 = R_{\mu\nu} = \frac{1}{2}Rg_{\mu\nu}$$

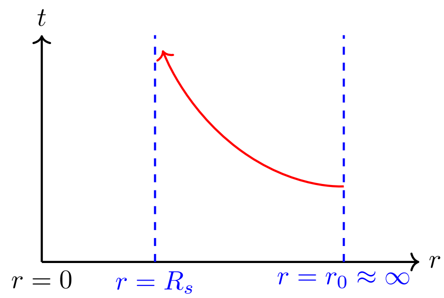
Does *not* indicate zero curvature $R_{\alpha\beta\mu\nu} = 0$. When studying the collapse of stars, there cases where the collapsing continues until electron degeneracy pressure supports the gravitational collapse of the star. These are **white dwarf** stars. If the original star is massive enough, gravity beats even electron degeneracy pressure and electrons are ejected and **neutron stars** are formed supported by neutron degeneracy pressure. Continuing this, even more massive stars can continue collapsing into super dense objects called **black holes**. For object with $r < R_s$ light can not escape (hence *black* holes).

In the early 1960s, GR became an important area of study and lots of work was done on black holes. At this time, people really began to consider black holes are actual physically real objects. By observing the orbits of stars in the center of the galaxy, black holes are indirectly measurable. [1]

The question becomes, what happens when a particle falls toward a black hole? Note that we have two singularities at $r = 0$ *and* $r = R_s$ (examine (6.9)). What happens to the curvature $R_{\alpha\beta\gamma\delta}$? Looking at the **Kretschmann scalar**,

$$K = R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} = \frac{12R_s^2}{r^6} \tag{6.10}$$

This is clearly divergent at $r = 0$, which indicates an infinite curvature. However at $r = R_s$, the curvature is well behaved.



Consider an observer at $r = \infty$ emitting a particle toward the black hole with decreasing r. Obviously, the original location of the observer $r = r_0 \approx \infty$ follows a geodesic. Furthermore, the timelike trajectory of the particle is also a geodesic. How much time is elapsed before the particle reaches $r = R_s$? This question is *ill-defined*. We need to measure time with respect to an observer. We will instead ask two questions,

- How much time elapses for the observer t?
- How much time elapses for the particle τ (proper time)?

To answer these questions, we will make use of the normalization condition (taking θ, ϕ constant),

$$-\left(1 - \frac{R_s}{r}\right)\dot{t}^2 + \left(1 - \frac{R_s}{r}\right)^{-1}\dot{r}^2 = -1$$

Here we have the conserved quantity,

$$E = -\left(1 - \frac{R_s}{r}\right)\dot{t} \tag{6.11}$$

Since there is no angular momentum (the particle is falling inward).

$$-\frac{E^2}{1 - \frac{R_s}{r}} + \frac{1}{1 - \frac{R_s}{r}}\dot{r}^2 = -1 \tag{6.12}$$

Solving this DE with initial conditions r_0, t_0, and $\dot{r}(r_0) = 0$ we can use the normalization condition to get,

$$-\frac{E^2}{1 - \frac{R_s}{r_0}} + 0 = -1$$

Which gives the conserved quantity E to be negative (since $\dot{t} > 0$ as time moves forward and $\frac{\mathrm{d}r}{\mathrm{d}t} < 0$ as the particle falls inward),

$$E = -\left(1 - \frac{R_s}{r_0}\right)^{1/2}$$

Taking the observer to be standing at $r_0 \approx \infty$, E where the observer stands is just,

$$E = -1$$

[1]To see a fancy video of this: https://www.youtube.com/watch?v=duoHtJpo4GY

Then (6.12) becomes,

$$-\frac{1}{1-\frac{R_s}{r}}+\frac{1}{1-\frac{R_s}{r}}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2=-1$$

Simplifying gives,

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2=\frac{R_s}{r}$$

Since r is decreasing with increasing τ,

$$\frac{\mathrm{d}r}{\mathrm{d}\tau}=-\left(\frac{R_s}{r}\right)^{1/2} \tag{6.13}$$

$$\Delta\tau=-\int_R^{R_s}\frac{r^{1/2}}{R_s^{1/2}}\mathrm{d}r=\frac{2}{3}\left(r^{3/2}\right)_{R_s}^{R}<\infty$$

Note this is just the proper time. So from the particle's perspective ($t_{\text{particle}}=\tau$) it reaches R_s in a finite time. What about Δt? The amount of time elapsed for the observer. How can we get rid of τ in (6.13)?

$$\frac{\mathrm{d}r}{\mathrm{d}\tau}=\frac{\mathrm{d}r}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau}=\frac{\mathrm{d}r}{\mathrm{d}t}\dot{t}$$

From the conserved quantity (6.11),

$$-1=-\left(1-\frac{R_s}{r}\right)\frac{\mathrm{d}t}{\mathrm{d}\tau}\implies\frac{\mathrm{d}\tau}{\mathrm{d}t}=\left(1-\frac{R_s}{r}\right)$$

Which gives,

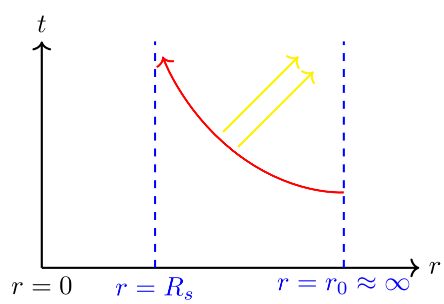
$$\frac{\mathrm{d}r}{\mathrm{d}t}=\frac{\mathrm{d}r}{\mathrm{d}\tau}\frac{\mathrm{d}\tau}{\mathrm{d}t}=-\left(1-\frac{R_s}{r}\right)\left(\frac{R_s}{r}\right)^{1/2}$$

Thus integrating gives,

$$\Delta t=-\int_R^{R_s}\frac{r}{R_s^{1/2}(1-\frac{R_s}{r})}\mathrm{d}r=+\infty$$

For an observer at ∞, it takes an ∞ amount of time to reach $r=R_s$.

As an extra question, we can ask what happens if the particle is emitting light back to the observer,



Obviously, the light emitted will be very red shifted,

$$\frac{\lambda_e}{\lambda_r}=\frac{\left(1-\frac{R_s}{r_e}\right)^{1/2}}{\left(1-\frac{R_s}{r_r}\right)^{1/2}}\sim_{r_r\to\infty}\left(1-\frac{R_s}{r_e}\right)^{1/2}\to 0$$

Therefore,

$$\lambda_r\gg\lambda_e$$

6.10 Singularities

Recall the Kretschmann Scalar given by (6.10). At $r = R_s$ the Kretschmann scalar and thus the scalar is finite. However for $r = 0$, $K \to \infty$ which is bad news. What is a true singularity? A singularity should be independent of coordinates. The problem at $r = R_s$ is just a manifestation of a poor choice of coordinates. As an example, consider $\mathbb{R}^3$. Suppose we are defining things in terms of spherical coordinates. If we have a function given by,

$$f(r, \theta, \phi) = \frac{1}{r-1}$$

This function f is perfectly well behaved everywhere except on the sphere with $r = 1$. This *could* in general mean that there is a problem with our manifold. However we know that Cartesian coordinates on $\mathbb{R}^3$ have no issues. It is always possible that divergence terms are a manifestation of a poor choice of coordinates.

Can we introduce another choice of coordinates so that the metric is well defined at $r = R_s$? As an exercise will will remove the $\mathrm{d}r^2$ term from the Minkowski metric.

Example: In the Minkowski case,

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2 \mathrm{d}\Omega^2$$

We will introduce a new variable $v = t + r$ such that,

$$\mathrm{d}v = \mathrm{d}t + \mathrm{d}r$$

Which gives the new metric,

$$\mathrm{d}s^2 = -\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r + r^2 \mathrm{d}\Omega^2$$

Here we have a new feature that this metric has non-diagonal terms. In a matrix representation,

$$g_{\alpha\beta} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & r^2 \end{bmatrix}$$

In order to explore the trajectory of light in this metric, we *could* look at the geodesic equation and compute the Christoffel symbols for this metric. More easily, looking at the normalization condition for a lightlike trajectory,

$$0 = -\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r + r^2 \mathrm{d}\Omega^2$$

Just focusing on radial motion, θ, ϕ are constants which means $\mathrm{d}\theta = \mathrm{d}\phi = 0$. This reduces the normalization condition to,

$$0 = -\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r$$

Factorizing,

$$0 = \mathrm{d}v(2\mathrm{d}r - \mathrm{d}v)$$

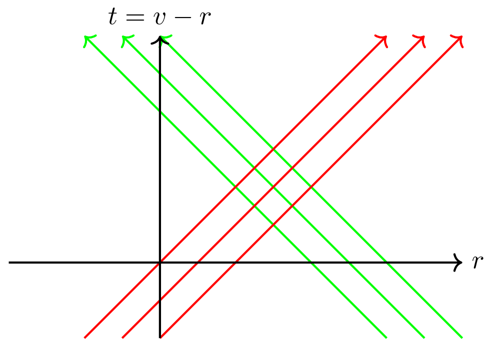
Which gives two possibilities. The first being that v is constant,

$$\mathrm{d}v = 0 \implies v = \text{const.} \implies t = \text{const} - r$$

These are drawn in green. Alternatively,

$$2\mathrm{d}r - \mathrm{d}v \implies t = r + \text{const}$$

These trajectories are drawn in red.



With the addition of radial symmetry, these trajectories or rays trace out light cones.

Example: Now in the case of the Schwarzschild metric,

$$v = t + r + R_s \ln\left|\frac{r}{R_s} - 1\right| \tag{6.14}$$

The metric then becomes,

$$\mathrm{d}s^2_{\text{sch}} \to \mathrm{d}s^2 = -\left(1 - \frac{R_s}{r}\right)\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r + r^2\mathrm{d}\Omega^2$$

This metric is known as the **Eddington-Finkenstein metric**. In a matrix representation,

$$g_{\alpha\beta} = \begin{bmatrix} -\left(1 - \frac{R_s}{r}\right) & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & r^2 \end{bmatrix}$$

It is important to note the subtlety here. Choosing spherical coordinates (t, r, θ, ϕ) resulted in a problem in the metric representation. However choosing better coordinates (t, v, θ, ϕ) results in the Eddington-Finkenstein metric. This metric has no divergence at $r = R_s$. The transformation from $v \to r$ diverges at $r = R_s$ (see (6.14)) because spherical coordinates were a bad choice, **not** because the manifold was a problem (this is contrasting to the case of $r = 0$ where there is no good choice of coordinates; a true singularity).

Now we can find the light trajectory by using the normalization condition and the purely radial motion $(\mathrm{d}\theta = \mathrm{d}\phi = 0)$.

$$-\left(1 - \frac{R_s}{r}\right)\mathrm{d}v^2 + 2\mathrm{d}v\mathrm{d}r = 0$$

Again factorizing,

$$\mathrm{d}v\left[-\left(1 - \frac{R_s}{r}\right)\mathrm{d}v + 2\mathrm{d}r\right] = 0$$

Which has two solutions,

$$\mathrm{d}v = 0 \implies v = \text{const}$$

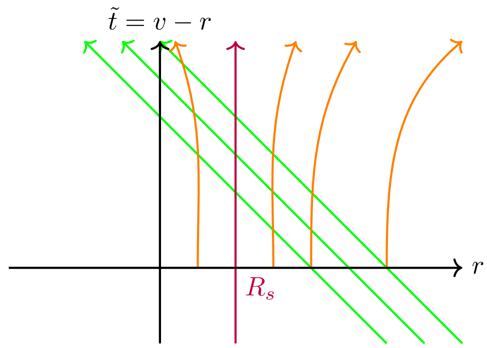
This case is drawn in green as before. Alternatively,

$$\left(1 - \frac{R_s}{r}\right)\mathrm{d}v = 2\mathrm{d}r$$

This differential equation has one solution at $r = R_s$ (purple). If $r \neq R_s$,

$$\mathrm{d}v = \frac{2\mathrm{d}r}{1 - \frac{R_s}{r}} \implies v = 2r + R_s \ln(r - R_s) + \text{const}$$

This solution is graphed below in orange.



The intersection between the orange and green trajectories moving forward (upward) in time form the line cones for the particles moving between them. For $r < R_s$, the trajectories for light fall inward toward $r = 0$. There is *no way* for light emanating from below $r < R_s$ to escape to a region where $r > R_s$. As such, for observers with $r > R_S$, $r = R_s$ acts as an **event horizon**. This is contrasting to the Minkowski case discussed above. An observer at any point in a Minkowski metric given enough time has access to information at every other point in spacetime.

7 Cosmology

7.1 Preface

Up until the notion of spacetime and GR, scientists assumed that the universe was infinite, static and eternal. Two characteristic assumptions of Cosmology are the notions of **homogeneity** and **isotropic**.

- homogeneous: Universe behaves/looks the same at every place
- isotropic: At every point, the universe looks the same in every direction

The first hint that suggests that the universe can't be static or infinite is that the night sky is mostly dark. If the universe were infinite, looking in any particular direction should be bright because *eventually* your ray of vision will hit a star. This idea was first discussed in 1826 and is known as **Albers' paradox**. Moreover in 1912, Slipher observed that galaxies are red shifted. This work was extended by Hubble in 1929 when he observed that this redshift is proportional to the distance to the galaxies. If galaxies are redshifted, then that means they are moving away from us. Combining this with the assumption of homogeneity, indicates that *all galaxies are moving away from each other*. Because of this, the universe is *not* static. Explicitly, this indicates that the space between galaxies are growing. In summary, these observations indicate that the universe can not be infinite or static.

At this time in history people started utilizing GR to study the *evolution* of the universe. To do so we will choose a metric to describe spacetime. Given the assumptions that *space alone* is homogeneous and isotropic we can arrive at some symmetries. Specifically the isotropic assumption introduces spherical symmetry and thus 3 rotation killing vectors. In additional homogeneity introduces "translational" invariance. Translational is written in quotes because it does not be linear translations in a Cartesian sense. As an example, consider a translation away from the north pole of a sphere. These "translationals" can manifest themselves as rotations or boosts. From this we also get 3 killing vectors.

7.2 Maximal Symmetry

Theorem: Given a manifold M with dimension $n > 1$, we can have at most $\frac{n(n+1)}{2}$ killing vectors. If we have the maximum number of killing vectors, we say that M is **maximally symmetric**. Moreover in the maximally symmetric case we have that,

$$R_{\alpha\beta\gamma\delta} = \frac{k}{n(n-1)}(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}) \tag{7.1}$$

Where k is just a constant with values potential values $-1, 0, 1$ and $g_{\alpha\beta}$ is a metric on M.

For example, taking space with 3 dimensions, the maximal number of killing vectors is,

$$\frac{n(n+1)}{2} = \frac{3(3+1)}{2} = 6$$

Therefore homogeneity and isotropy immediately imply that we have a maximally symmetric space. One example is hyperplane $\mathbb{R}^3$ with zero curvature $k = 0$. A second example is S^3 the 3-sphere with curvature $k = 1$ and finally the Hyperboloid with curvature $k = -1$.

Space	Curvature k
$\mathbb{R}^3$	0
S^3	+1
$\mathcal{H}^3$	-1

7.3 FLRW Metric

In order to obtain the spatial metric, we can solve (7.1). A general expression for a metric can be written with $i = 1, 2, 3$,

$$g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta = g_{00}(\mathrm{d}t)^2 + g_{0i}\mathrm{d}t\mathrm{d}x^i + g_{i0}\mathrm{d}x^i\mathrm{d}t + g_{ij}(\vec{x}, t)\mathrm{d}x^i\mathrm{d}x^j \tag{7.2}$$

First we take a look at the space components of the metric γ_{ij} and utilize the fact that space is homogeneous and isotropic metric,

$$\mathrm{d}\sigma^2 = \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = e^{2\beta(r)}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \tag{7.3}$$

We can using (7.1) to solve for $\beta(r)$. Computing the Ricci tensor from (7.1) gives,

$$R_{\beta\delta} = R_{\alpha\beta\delta\gamma}g^{\alpha\gamma} = \frac{k}{6}\left(\underbrace{g_{\alpha\gamma}g^{\alpha\gamma}}_{3} g_{\beta\delta} - \underbrace{g^{\alpha\gamma}g_{\alpha\delta}}_{\delta^\gamma{}_\delta} g_{\beta\gamma}\right)$$

$$R_{\beta\delta} = \frac{k}{6}(2g_{\beta\delta})$$

Therefore (7.3) becomes,

$$\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = \frac{1}{1-kr^2}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \tag{7.4}$$

Now we can return to our general metric (7.2) and analyze the time components. To do so we will use the notion of **comoving coordinates**. Here the tangent vector of a world line is always perpendicular to hypersurface of space and naturally the position components are held constant $x^i = \text{const}$ (Wainwright's notes p114). As such we can make the off diagonal terms zero,

$$g_{0i} = g_{i0} = 0$$

And the first components can be negative,

$$g_{00} = -1$$

What about g_{ij} in (7.2)? If in general $g_{ij} = g_{ij}(\vec{x}, t)$ the space metric depends on $\vec{x}$ then we violate the assumption of homogeneity. Therefore it must depend on r. Similarly, if g_{ij} is to be isotropic in time, the time components must be separable from g_{ij}. Recalling (7.4),

$$g_{ij}(\vec{x}, t) \to a^2(t)\gamma_{ij}(r)$$

Here $a(t)$ is known as a **scale factor**. Moreover the square on $a(t)$ is introduced to ensure positiveness. Combining all of these motivations together yields the **Friedman Le maître Robertson Walker metric (FLRW).**

$$g_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta = -\mathrm{d}t^2 + a^2(t)\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{7.5}$$

The discovery and proof of uniqueness of this metric took from ∼1922-1935. What are the implications of this metric? What we will see is that (7.5) produces a relationship between redshift and distance.

7.4 Redshift & Distance

Suppose we have two individuals each in a separate galaxy very far apart. If person 1 sends a signal to person 2 at a time $t = t_e$ at some later time, person B will receive this signal. If their separation was initially ℓ, by (7.5), the separation at a later time t_r is given by $a(t_r)\ell$. Denote the tangent vectors for each observed V_1^μ and V_2^μ respectively. Finally, let the receiver of light (2) be placed at $r = 0$. Since the light signal is moving along a geodesic (and the motion is purely radial), the normalization condition for a lightlike trajectory becomes,

$$0 = -\dot{t}^2 + a^2(t)\frac{\dot{r}^2}{1 - kr^2}$$

Therefore since the position of the light relative to observer (2) r is decreasing, $\mathrm{d}r < 0$. By factoring out the $\mathrm{d}\tau$'s one obtains,

$$\frac{\mathrm{d}t}{a(t)} = -\frac{\mathrm{d}r}{(1 - kr^2)^{1/2}}$$

Integrating over the path of the trajectory,

$$\int_{t_e}^{t_r} \frac{\mathrm{d}t}{a(t)} = -\int_{r_e}^{r_r} \frac{\mathrm{d}r}{(1 - kr^2)^{1/2}} = f_k(r_r, r_e) \tag{7.6}$$

We will not consider a series of light pulses spaced by short intervals δt_e with frequency $\omega_e = \frac{2\pi}{\delta t_e}$ when emitted and received with frequency $\omega_r = \frac{2\pi}{\delta t_r}$. For every pulse, (7.6) holds true. For the pulse at $t = t_e + \delta t_e$ we can assume that δt_e is small enough or that $a(t)$ is small enough such that the total *distance* traveled does not differ from the original pulse,

$$\int_{t_e+\delta t_e}^{t_r+\delta t_r} \frac{\mathrm{d}t}{a(t)} = -\int_{r_e}^{r_r} \frac{\mathrm{d}r}{(1 - kr^2)^{1/2}} = f_k(r_r, r_e)$$

Since the traveled distance is the same we have,

$$\int_{t_e+\delta t_e}^{t_r+\delta t_r} \frac{\mathrm{d}t}{a(t)} = \int_{t_e}^{t_r} \frac{\mathrm{d}t}{a(t)}$$

Fragmenting the bounds of integration,

$$\int_{t_e+\delta t_e}^{t_r+\delta t_r} \frac{\mathrm{d}t}{a(t)} = \int_{t_r}^{t_r+\delta t_r} \frac{\mathrm{d}t}{a(t)} + \int_{t_e}^{t_r} \frac{\mathrm{d}t}{a(t)} + \int_{t_e+\delta t_e}^{t_e} \frac{\mathrm{d}t}{a(t)}$$

The term over $t_e \to t_r$ cancels out giving,

$$\int_{t_e}^{t_e+\delta t_e} \frac{\mathrm{d}t}{a(t)} = \int_{t_r}^{t_r+\delta t_r} \frac{\mathrm{d}t}{a(t)}$$

Finally using the fundamental theorem of calculus $(F(t + \delta t) - F(t) \approx \delta t F'(t))$,

$$\frac{\delta t_r}{a(t_r)} = \frac{\delta t_e}{a(t_e)}$$

Rearranging and switching to frequencies and wavelengths gives the redshift factor Z,

$$\frac{\delta t_r}{\delta t_e} = \frac{\omega_e}{\omega_r} = \frac{\lambda_r}{\lambda_e} = \frac{a(t_r)}{a(t_e)} \equiv 1 + Z$$

So does (7.5) produce a redshift or a blueshift? This depends on the change of a. If $a(t)$ is decreasing, we will see a blueshift of distance galaxies. Conversely if $a(t)$ is increasing, we should expect to observe a redshift.

Using this result, what can we say about the relationship between redshift and distance? Consider the Taylor expansion of $a(t)$,

$$a(t) = a(t_0) + (t - t_0)\dot{a}(t_0) + \frac{1}{2}(t - t_0)^2\ddot{a}(t_0) + O\Big((t - t_0)^3\Big) \tag{7.7}$$

We will now define some parameters. The **Hubble parameter** is,

$$H(t) = \frac{\dot{a}(t)}{a(t)}$$

The **deceleration parameter** is,

$$q(t) = -\frac{a(t)\ddot{a}(t)}{\dot{a}^2(t)}$$

Of course, the corresponding higher order terms are called **Jerk, Snap, Crackle** and **Pop** respectively. Moving forward we will look at the first order terms in (7.7) and treat t and the time emitted and t_0 the time received,

$$a(t_e) = a(t_r) + (t_e - t_r)\dot{a}(t_r)$$

Rearranging yields,

$$\frac{a(t_e) - a(t_r)}{a(t_r)} = (t_e - t_r)\frac{\dot{a}(t_r)}{a(t_r)} = (t_e - t_r)H(t_r)$$

Where H is just the Hubble parameter,

$$\frac{a(t_e)}{a(t_r)} - 1 = H(t_r)(t_e - t_r)$$

Recalling the definition of redshift $1 + Z = \frac{a(t_r)}{a(t_e)}$,

$$\frac{1}{1 + Z} - 1 = H(t_r)(t_e - t_r)$$

For small $Z \ll 1$,

$$-Z \approx H(t_r)(t_e - t_r)$$

The distance that the light traveled is given by $R = a(t_r)\ell$. But,

$$t_r - t_e = \Delta t = \frac{R}{c} = R \quad \text{With } c = 1$$

Giving a linear relation between redshift and distance,

$$Z = H(t_r)R = H_0 R$$

Of course this only works for small redshift; or for object a few megaparsecs away. This linear result between distance R and Z was observed by Hubble himself using variable stars specifically Cepheids. The constant H_0 that Hubble observed in 1929 was,

$$H_0 = 500\,\mathrm{kms}^{-1}/\mathrm{Mpc}$$

Where as in the 2000s, this measurement has been refined to,

$$H_0 = 67\,\mathrm{kms}^{-1}/\mathrm{Mpc}$$

7.5 Stress Energy Tensor of Perfect Fluids

Recall that in General relativity we have the Einstein equation that indicates that stuff tells how gravity should be. We also have geodesics which indicate how particles propagate. Finally, we have the conservation of the Stress Energy Tensor which indicates how fluids behave. We can then assume that the stuff in the universe behaves like a perfect fluid that is completely characterized by (ρ, P). Perfect fluids have the equation of states,

$$P = w\rho$$

Where w characterizes the type of fluid. For some examples,

Dust (Matter): Dust have no pressure. $P = 0$ gives us that $w = 0$.

Radiation: The stress energy tensor is given by,

$$T^{\text{EM}}_{\mu\nu} = F_{\mu\alpha}F_{\nu}{}^{\alpha} - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}$$

Where $F_{\alpha\beta}$ is the Maxwell stress tensor and the trace of $T^{\text{EM}}{}_{\mu\nu}$ is zero,

$$g^{\mu\nu}T^{\text{EM}}_{\mu\nu} = F_{\mu\alpha}F^{\mu\alpha} - \frac{1}{4}\underbrace{g_{\mu\nu}g^{\mu\nu}}_{4}F_{\alpha\beta}F^{\alpha\beta} = 0$$

If radiation is approximated by a perfect fluid we should have a traceless stress energy tensor,

$$T^{\alpha\beta} = (\rho + P)V^{\alpha}V^{\beta} + Pg^{\alpha\beta}$$

We can use co-moving coordinates, $V^{\alpha} = (1, 0, 0, 0)$,

$$T^{\alpha\beta} = \begin{pmatrix} \rho & 0 \\ 0 & g^{ij}P \end{pmatrix}$$

The trace then is,

$$T^{\alpha\beta}g_{\alpha\beta} = -\rho + \underbrace{g_{ij}g^{ij}}_{3}P = 0$$

Therefore,

$$P = \frac{1}{3}\rho$$

Which gives that our w term must be,

$$w = \frac{1}{3}$$

For a perfect fluid interpreted as radiation.

Cosmological Constant: Using the FLRW metric and looking at the Einstein equations,

$$G_{\alpha\beta} = kT_{\alpha\beta}$$

With $k = 8\pi G$. Expanding out $G_{\alpha\beta}$,

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} + \Lambda g_{\alpha\beta} = kT_{\alpha\beta}$$

We can take the cosmological constant term and move it to the RHS,

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = kT_{\alpha\beta} - \Lambda g_{\alpha\beta} = k\big(T_{\alpha\beta} - T^{\Lambda}_{\alpha\beta}\big)$$

Where,

$$T^{\Lambda}_{\alpha\beta} = \begin{pmatrix} \frac{\Lambda}{k} & 0 \\ 0 & -\frac{\Lambda}{k}g_{ij} \end{pmatrix}$$

Where $\rho^{\Lambda} = \frac{\Lambda}{k}$ and $P^{\Lambda} = -\frac{\Lambda}{k}$. This gives,

$$w = -1$$

When the Cosmological constant is interpreted as a perfect fluid.

In summary we have,

Material	w
Dust (Matter)	0
Radiation	$\frac{1}{3}$
Cosmo. Const.	-1
"Curvature"	$-\frac{1}{3}$

We will now look at the equation of motion of the fluid; namely the conservation of the stress energy tensor,

$$\nabla^{\mu} T_{\mu\alpha} = 0$$

Again using the FLRW metric,

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j$$

In order to examine the conservation equation, we must first compute the associated Christoffel symbols. In general it follows in a typical way,

$$\Gamma^0{}_{11} = \frac{a\dot{a}}{1-kr^2} \cdots$$

One result that stems from computing the Christoffel symbols yields,

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P) \tag{7.8}$$

Recall that the form of gravity is completely characterized by the scale factor a. Plugging in the equation of state $P = w\rho$ gives,

$$\dot{\rho} = -3\frac{\dot{a}}{a}\rho(1+w)$$

Obviously depending on the type of material considered (w) we can solve this differential equation for ρ in terms of a.

$$\rho = \rho_0 a^{-3(1+w)}$$

Therefore subbing in different materials,

Material	w	ρ/ρ_0
Dust (Matter)	0	a^{-4}
Radiation	$\frac{1}{3}$	a^{-3}
Cosmo. Const.	-1	1
"Curvature"	$-\frac{1}{3}$	a^{-2}

Since we observe redshift in the universe, it must be that a is increasing with time. Clearly, the power on a indicates which materials are dominate in the Universe. When a is small, radiation is dominant. As a grows, the dominant term becomes matter and then finally for a very large we obtain that the cosmological is dominant.

Note: An easier way to determine (7.8) is to utilize the result,

$$\nabla_{\alpha} V^{\alpha} = |\det g|^{-\frac{1}{2}} \partial_{\alpha}\left(|\det g|^{\frac{1}{2}} V^{\alpha}\right) \tag{7.9}$$

Where $\det g$ is the determinant of the metric $g_{\mu\nu}$. Before utilizing this result, we first have to derive a relativistic continuity equation for perfect fluids. Beginning with the stress energy tensor,

$$T^{\mu\nu} = (\rho + p)V^{\mu}V^{\nu} + pg^{\mu\nu}$$

Taking V^μ to be a timelike tangent to a geodesic and contracting the SET with $V_\nu \nabla_\mu$,

$$V_\nu \nabla_\mu (T^{\mu\nu}) = V_\nu \nabla_\mu ((\rho + p) V^\mu V^\nu + p g^{\mu\nu})$$

Obviously, the conservation of the SET gives $\nabla_\mu (T^{\mu\nu}) = 0$. Expanding out the terms on the RHS,

$$0 = V_\nu \nabla_\mu ((\rho + p) V^\mu V^\nu) + V_\nu \nabla_\mu (p g^{\mu\nu})$$

Using product rule,

$$0 = V_\nu \nabla_\mu (\rho + p) V^\mu V^\nu + V_\nu (\rho + p) \nabla_\mu (V^\mu) V^\nu + V_\nu (\rho + p) V^\mu \nabla_\mu (V^\nu) + V_\nu \nabla_\mu (p g^{\mu\nu})$$

Now since $V^\alpha V_\alpha = -1$,

$$0 = -V^\mu \nabla_\mu (\rho + p) - (\rho + p) \nabla_\mu (V^\mu) + (\rho + p) V^\nu V_\mu \nabla_\mu (V^\nu) + V_\nu \nabla_\mu (p g^{\mu\nu})$$

Now since V^α is a geodesic, $V_\nu \nabla_\mu (V^\nu) = 0$. Also utilize that the connection is metric compatible,

$$0 = -V^\mu \nabla_\mu (\rho + p) - (\rho + p) \nabla_\mu (V^\mu) + V_\nu \nabla^\nu p$$

The pressure terms cancel leaving,

$$0 = V^\mu \nabla_\mu \rho + (\rho + p) \nabla_\mu (V^\mu)$$

Also note that ρ is a scalar field so by definition,

$$0 = V^\mu \partial_\mu \rho + (\rho + p) \nabla_\mu (V^\mu)$$

Since th FRW metric has $\det g = (-1)\left(a^2\right)\left(a^2\right)\left(a^2\right) = -a^6$, (7.9) reduces to,

$$\nabla_\mu (V^\mu) = 3\frac{\dot{a}}{a}$$

Provided that comoving coordinates are chosen $V^\alpha = (1, 0, 0, 0)$. Then we have that,

$$\dot{\rho} + 3\frac{\dot{a}}{a}(p + \rho) = 0$$

Which is precisely (7.8).

7.6 Einstein Equations With Stress Energy Tensor

So how does a change with time explicitly? To discover this we consider the Einstein equation where we introduce the different densities,

$$\rho^{\text{tot}} = \rho_M + \rho_R + \rho_\Lambda + \rho_C$$

Where the time evolution is characterized by,

$$\rho_I(t) = \rho_{0I} a^{-m_I}$$

Where for each material I,

$$m_I = (4, 3, 2, 0)$$

The Einstein equations are then,

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T^{\text{tot}}_{\mu\nu}$$

We will use the FLRW metric to calculate $R_{\mu\nu}$ for this metric. This is left as an exercise. As an example,

$$R_{00} = -3\frac{\ddot{a}}{a}$$

Then we obtain that (with $k = 1, 0, -1$ according to the curvature of space) in the comoving frame,

$$G_{00} = \frac{3}{a^2}\left(k + \dot{a}^2\right) = 8\pi G\rho = 8\pi G(\rho_R + \rho_M + \rho_\Lambda) \tag{7.10}$$

And the space terms,

$$G_{ii} = -\left(2\frac{\ddot{a}}{a} + \frac{1}{a^2}\left(k + \dot{a}^2\right)\right) = 8\pi G\rho \tag{7.11}$$

Using both (7.10) and (7.11) to solve for $\dot{a}$ and $\ddot{a}$,

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - k \tag{7.12}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{7.13}$$

These equation we first obtained by Friedman in 1922 and later by Le maître 1927. Even without solving these equations we can learn a lot by making astronomical observations. We can learn about,

1. The nature of space (i.e. the value of k)
2. The age of the Universe (approximately)
3. The fate of the Universe

7.6.1 The Nature of Space

The learn about the age of the universe, we will take (7.12) and divide by a^2. As you will see, we obtain the Hubble constant H,

$$H^2 = \frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\rho - \frac{k}{a^2}$$

Dividing by H^2,

$$1 = \frac{8\pi G}{3H^2}\rho - \frac{k}{H^2 a^2}$$

More rearranging,

$$\frac{k}{H^2 a^2} = \frac{8\pi G}{3H^2}\rho - 1$$

Defining $\rho_{\text{crit}}(t) = \frac{3H^2(t)}{8\pi G}$,

$$\frac{k}{H^2 a^2} = \frac{\rho}{\rho_{\text{crit}}} - 1$$

Taking $t = t_{\text{now}}$ the time it is now we can then measure $H_0 = H(t_{\text{now}})$ via redshift. We can use this to measure $\rho_{\text{crit}}(t_{\text{now}})$. From this can measure ρ and its components (ρ_M, ρ_R).

From observations we measure ρ and get the value of $\rho/\rho_{\text{crit}} = \Omega$,

$$\frac{k}{H^2 a^2} = \Omega - 1$$

Notice that $H^2 a^2$ is positive. Therefore the value of Ω determines the sign of k and thus the nature of space! If from observations we obtain $\rho > \rho_{\text{crit}}$, we get that $\Omega > 1$ and then $k > 0$ and thus space is a sphere. If $\rho < \rho_{\text{crit}}$, $k < 0$ and we live in a hyperboloid. Finally if $\rho = \rho_{\text{crit}}$, we get $\Omega = 1$ and thus $k = 0$ and we live in a hyperplane $\mathbb{R}^2$. Our current observation is that,

$$\rho = \rho_{\text{crit}}$$

What we will see is that the value of k determines the fate of the Universe.

7.6.2 Age of The Universe

The age of the Universe comes from (7.13).

$$3\frac{\ddot{a}}{a} = -(4\pi G(\rho + 3P) - \Lambda)$$

Looking more carefully at the inner term $\rho + 3P$. $\rho + 3P$ is positive if $P > -\frac{1}{3}\rho$ or if $w < -\frac{1}{3}$. Considering radiation and matter, $\rho + 3P > 0$. Moreover looking at Λ, if $\Lambda < 0$ the overall RHS is negative meaning $\ddot{a} < 0$; the Universe is decelerating. However if $\Lambda > 0$ the Λ contribution is causing an acceleration of the Universe (of course the mass/radiation term always caused deceleration). From current observation suggests that $\Lambda \approx 0^+$ is slightly bigger than zero. For simplicity we sat that $\Lambda = 0$. Therefore $\ddot{a} < 0$. But from observation we have that $\dot{a} > 0$ since we observe an expanding Universe. Therefore the curve of a *must* be concave down. How a behaves for large t has not been answered yet but right now, $a(t \approx t_{\text{now}})$ is concave down. Taking a tangent line along a backward in time,

$$a_{\tan} - a(t_{\text{now}}) = \dot{a}(t_{\text{now}})(t - t_{\text{now}})$$

Obviously when $a = 0$ we have $a(t_{\text{bang}}) = 0$ a time when there was a t_{bang}. We will call the approximation using the tangent line to be $a_{\tan}(t_{\text{boom}}) = 0$.

$$-\frac{a(t_{\text{now}})}{\dot{a}(t_{\text{now}})} = (t_{\text{boom}} - t_{\text{now}})$$

$$\frac{1}{H_0} = t_{\text{now}} - t_{\text{boom}}$$

But $t_{\text{now}} - t_{\text{boom}} < t_{\text{now}} - t_{\text{bang}}$ Therefore,

$$\frac{1}{H_0} > t_{\text{now}} - t_{\text{bang}} = \text{Age of Universe}$$

7.7 Fate of The Universe

The future of the universe is determined by (7.12),

$$\dot{a}^2 = \frac{8\pi}{3}G\rho a^2 - k$$

Where as we have seen before with $w > -1/3$,

$$\rho a^2 = a^{-3(1+w)+2}$$

Where when $a \to \infty$, $\rho a^2 \to 0$. Then for different values of k we can see that the fate of the universe will be.

Forever Expansion $\mathbf{k = -1}$: If $k = -1$, then by observation of $\dot{a} > 0$ with $\dot{a} \neq 0$. This corresponds to expansion forever.

Slowing Expansion $\mathbf{k = 0}$: If $k = 0$ we have that $\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 \to 0$ whenever a is large. Therefore we still have an expansion forever but the expansion becomes slower and slower with time.

Eventual Contraction $\mathbf{k = 1}$: If $k = 1$, we have $\dot{a}^2 = \frac{8\pi}{3}G\rho a^2 - k$ achieves a zero value at a critical a. By observation again $\dot{a} > 0$, there exists a particular time $t_{\max}$ when,

$$a(t_{\max}) = a_{\max} \implies \dot{a}(t_{\max}) = 0$$

For $t > t_{\max}$ we will reach a point where $\dot{a} < 0$. In this case the universe starts contracting until we reach a big crunch.

As argued before, we can assume that the space curvature contribution k is in fact the contribution of a fluid with equation of state $P = -\frac{1}{3}\rho$. Writing all three contributions,

$$\dot{a}^2 = \frac{8\pi G}{3}\sum_{b=R,M} c_b a^{-(1+3w_b)} - k + \frac{\Lambda}{3}a^2$$

Which becomes,

$$\dot{a}^2 = \frac{c_M}{a} + \frac{c_R}{a^2} - k + \frac{\Lambda}{3}a^2 \tag{7.14}$$

According to the value of a at different times in the evolution of the universe, different contributions will dominate.

As time evolves, we have the dominating term evolve,

$$\text{small } a : \text{radiation} \to \text{matter} \to \text{curvature} \to \text{cosmological constant} : \text{large } a$$

7.8 Einstein Universe

In the olden times before Hubble, the Universe was thought to be static with $\dot{a}(t) = 0 \to a(t) = a_0, \forall t$. Using this idea in the conservation of the SET and in the Friedman equations respectively,

$$\dot{\rho} = -3(\rho + p)\frac{\dot{a}}{a} \implies \dot{\rho} = 0 \implies \rho = \rho_0 \tag{7.15}$$

$$-3\frac{\ddot{a}}{a} = 4\pi G(\rho + 3p) - \Lambda \implies \Lambda = 4\pi G(\rho + 3p) \tag{7.16}$$

When considering matter or radiation, equation (7.16) requires that $\Lambda > 0$. This was the source of Einstein's motivation to introduce Λ in the first place; to make the universe static. We also obtain that $\dot{\rho} = 0$ from (7.15) which should make sense for a static universe.

Next we will need to determine what type of spatial geometry we are dealing with. First looking at the Friedman equation,

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi}{3}G\rho + \frac{\Lambda}{3}$$

Which for a static universe gives.

$$\frac{k}{a_0^2} = 4\pi G(\rho + p) > 0$$

If $k > 0$, then it must be that $k = 1$ and,

$$a_0^2 = (4\pi G(\rho + p))^{-1}$$

Recall from section 7.2 that $k = 1$ corresponds to space being a *sphere* (topological $\mathbb{R} \times S^3$). This is evident because as the effective radius scale a_0 decreases, the density is increasing. Therefore the cosmological constant exactly matches the gravitational attraction.

$$\ddot{a}_0 = -\frac{1}{3}(4\pi G(\rho + 3p) - \Lambda)a_0 = 0$$

Since $\ddot{a}_0 = 0$, it is unstable under small perturbations $a = a_0 + \epsilon a_1$. A Universe with only matter has $(\Lambda = 0, k = 0)$. This is known as **Einstein de Sitter Universe**. In this universe we have to solve (7.14) and look at the case of $k = 0$.

$$(\dot{a})^2 = \frac{c_M}{a} \implies \dot{a} = \pm\left(\frac{c_M}{a}\right)^{1/2}$$

In the $\dot{a} > 0$ case,

$$a(t) \propto (t)^{2/3}$$

This is the Einstein de Sitter universe where the universe is in expansion from the Big bang onward when $t = 0$. In the $\dot{a} < 0$ case,

$$a(t) \propto (t_f - t)^{2/3}$$

Which corresponds to a collapsing universe with a Big Crunch at $t = t_f$. Alternatively, one can examine a universe with just a cosmological constant (no matter/radiation).

$$\dot{a}^2 = -k + \frac{\Lambda}{3} a^2$$

Since $\dot{a}^2 > 0$ always, if $\Lambda > 0$ then $k = -1, 0, 1$ are all allowed. However when $\Lambda < 0$, it is required that $k = -1$. A new notation that is common to adopt is $\left|\frac{\Lambda}{3}\right| = \pm \ell^{-2}$. Then we have.

$$\dot{a}^2 = -k \pm a^2 \ell^{-2} \implies \ddot{a} = \pm a \ell^{-2}$$

Therefore there are 4 cases to consider.

$\boldsymbol{\Lambda > 0}$ Metrics of the de Sitter spacetime:

$$\begin{aligned} k = 0 &\to a_+(t) = e^{+\frac{t}{\ell}} \to \mathrm{d}\tau^2 = -\mathrm{d}t^2 + e^{+\frac{t}{\ell}} \mathrm{d}\vec{x}^2 \\ k = 1 &\to \mathrm{d}\tau^2 = -\mathrm{d}t^2 + \ell^2 \cosh^2\left(\frac{t}{\ell}\right) \mathrm{d}\tilde{\Omega}^2 \\ k = -1 &\to \mathrm{d}\tau^2 = -\mathrm{d}t^2 + \ell^2 \sinh^2\left(\frac{t}{\ell}\right) \mathrm{d}\tilde{\Omega}^2 \end{aligned}$$

$\boldsymbol{\Lambda < 0}$ Metrics of the Anti de Sitter spacetime:

$$k = -1 \to \mathrm{d}\tau^2 = -\mathrm{d}t^2 + \ell^2 \sin^2\left(\frac{t}{\ell}\right) \mathrm{d}\tilde{\Omega}^2$$

8 Gravitational Waves

Note by Author: This section was delivered as a set of slides in the last part of the last lecture. The aim of these section to shed light on some of the techniques of perturbation theory in context of general relativity. The logical flow is not meant to be flawless as it is merely a summary of the presented mathematics. *It is not considered testable material.*

8.1 Perturbations

In physics, we often analyze a system by looking at what happens under small perturbations (cf quantum field theory). To simplify the Einstein equations, we can consider a *fixed background metric* $g^{(0)}_{\mu\nu}$ and look at the behavior of *perturbations* $h_{\mu\nu}$ around it,

$$g_{\mu\nu} = g^{(0)}_{\mu\nu} + h_{\mu\nu} \qquad |h_{\mu\nu}| \ll 1$$

To further simplify the analysis, we can consider the background metric to be the Minkowski metric with signature $(-, +, +, +)$.

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

Here $h_{\mu\nu}$ is the metric perturbation. In effect we are going to study the dynamics of the symmetric tensor field $h_{\mu\nu}$ propagating on Minkowski spacetime. This is a *linearized General Relativity*. In order to look at the Einstein equations, we first need to known the inverse metric, Christoffel symbols and the Riemann tensor. Keeping only **the first order terms**, the inverse metric is,

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

The Christoffel symbols are determined by the metric ($\partial_\alpha \eta_{\beta\gamma} = 0$),

$$\Gamma^\rho{}_{\mu\nu} = \frac{1}{2} g^{\rho\lambda}(-\partial_\lambda g_{\mu\nu} + \partial_\mu g_{\nu\lambda} + \partial_\nu g_{\lambda\mu}) = \frac{1}{2}\eta^{\rho\lambda}(-\partial_\lambda h_{\mu\nu} + \partial_\mu h_{\nu\lambda} + \partial_\nu h_{\lambda\mu})$$

And the Riemann tensor is,

$$R_{\mu\nu\rho\sigma} = \eta_{\mu\lambda}\partial_\rho \Gamma^\lambda{}_{\mu\sigma} - \eta_{\mu\lambda}\partial_\sigma \Gamma^\lambda{}_{\mu\rho} = \frac{1}{2}(\partial_\rho\partial_\nu h_{\mu\sigma} + \partial_\sigma\partial_\mu h_{\nu\rho} - \partial_\sigma\partial_\nu h_{\mu\rho} - \partial_\rho\partial_\mu h_{\nu\sigma})$$

Using the Riemann tensor, we can obtain the Ricci tensor and Ricci scalar,

$$R_{\mu\nu} = \frac{1}{2}(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\nu - \partial_\mu\partial_\nu h - \Box h_{\mu\nu})$$

Where h is just the trace of the metric perturbation $h = \eta^{\mu\nu}h_{\mu\nu}$ and $\Box$ is the Alembertian. Then one can obtain the Ricci scalar,

$$R = \partial_\mu\partial_\nu h^{\mu\nu} - \Box h$$

The Einstein tensor becomes,

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}R$$

$$G_{\mu\nu} = \frac{1}{2}(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\nu - \partial_\mu\partial_\nu h - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial_\sigma\partial_\rho h^{\sigma\rho} + \eta_{\mu\nu}\Box h) \tag{8.1}$$

Which gives a corresponding equation of motion,

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}$$

Or in a vacuum,

$$G_{\mu\nu} = 0$$

8.2 Modes of Propagation

Now we can consider the linearized Einstein equations term by term in order to determine the propagating modes. The modes will be denoted,

Mode Type	Notation
Scalar	Φ
Vector	w_i
Tensor	s_{ij}

To clarify the analysis, we can write each of the perturbation elements explicitly,

$$h_{00} = -2\Phi \qquad h_{0i} = h_{i0} = w_i \qquad h_{ij} = 2s_{ij} - 2\delta_{ij}\mathfrak{s}$$

Where $\mathfrak{s}$ is proportional to the trace of the tensor mode,

$$\mathfrak{s} = -\frac{1}{6}s_i{}^i = -\frac{1}{6}s_{ij}\delta^{ij}$$

The tensor mode is traceless and called the strain,

$$s_{ij} = \frac{1}{2}\left(h_{ij} - \frac{1}{3}\delta_{ij}h_{kl}\delta^{kl}\right)$$

The scalar mode is related to the Newton potential.

The Einstein equations become 3 sets of equations for each of the modes,

$$G_{00} = s\nabla^2\mathfrak{s} + \partial_k\partial_l s^{kl} = 8\pi G T_{00}$$

This determines $\mathfrak{s}$ in terms of the strain , the source and the boundary conditions. It is not propagating.

$$G_{0j} = -\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j\partial_k w^k + 2\partial_0\partial_j\mathfrak{s} + \partial_0\partial_k s^k{}_j = 8\pi G T_{0j}$$

This determines the vector mode in terms of the strain, the source and the boundary conditions. It is not propagating. And finally,

$$G_{ij} = \left(\delta_{ij}\nabla^2 - \partial_i\partial_j\right)(\Phi - \mathfrak{s}) + \delta_{ij}\partial_0\partial_k w^k - \partial_0\left(\partial_{(i}w_{j)}\right) + 2\delta_{ij}\partial_0^2\mathfrak{s} - \Box s_{ij} + 2\partial_l\left(\partial_{(i}s_{j)}{}^k\right) - \delta_{ij}\delta_k\delta_l s^{kl} = 8\pi G T_{ij}$$

This determines the scalar mode in terms of the strain, the source and the boundary conditions. It is not propagating. The strain is therefore the only propagating degrees of freedom.

8.3 Diffeomorphisms

We have only few dynamical variables. We can use the symmetries of the Einstein equations to eliminate the spurious variables and focus on the relevant ones, just as in electromagnetism. Diffeomorphisms, or change of coordinates, provide the symmetry of the Einstein equations.

$$x^\mu \to x'^\mu \qquad g_{\mu\nu}(x) \to g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x)$$

The understanding of the diffeomorphism symmetry took a long time for Einstein to realize this was a key element in the construction of GR (cf hole argument in Rovelli's book "Quantum Gravity"). Still now, this is a topic of questioning (e.g. what is time?) and often confusing. The bottom line is that:

- the observable quantities should be invariant under diffeomorphisms
- the coordinates used to parameterize GR are not physical, but it is possible to introduce some (idealized) degrees of freedom as clocks and rulers to define physical coordinates (cf the GPS). A typical way to introduce these (idealized) degrees of freedom is to introduce a gauge (just as in electromagnetism).

8.4 Gauge

The general diffeomorphisms will not preserve the linearized Einstein equations. We therefore look at their linearized version.

$$x^\mu = x^\mu + \xi^\mu(x) \qquad h'_{\mu\nu}(x) = h_{\mu\nu} - \partial_\mu\xi_\nu - \partial_\nu\xi_\mu$$

Let us come back to the linearized Einstein equations (8.1) and perform a gauge transformation (i.e. choose a coordinate system) such that the equations become simpler. In particular the **Lorentz Gauge**,

$$\partial_\sigma h^\sigma{}_\mu - \frac{1}{2}\partial_\mu h = 0$$

Makes (8.1) become,

$$-\Box h_{\mu\nu} + \frac{1}{2}\eta_{\mu\nu}\Box h = 4\pi G T_{\mu\nu}$$

In fact we have the possibility to simplify further the metric perturbations by going to the radiation gauge, which freezes the trace and the scalar and vector modes to zero when we have no source (cf Wald p.80). The gauge is also called the Transverse-Traceless (TT) gauge.

$$\Box h_{\mu\nu} = 0$$

This is just the wave equation for $h_{\mu\nu}$.

$$h_{\mu\nu} = a_{\mu\nu}e^{ix_\mu k^\mu}$$

Where $a_{\mu\nu}$ is just a constant with,

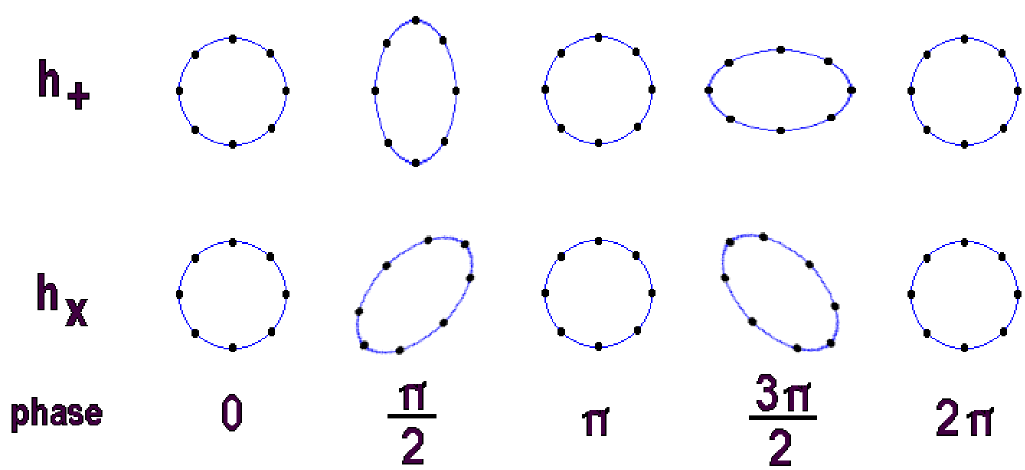
$$a_{\mu\nu} = \begin{cases} a_{0\mu} = 0 & \text{radiation gauge} \\ a_\mu{}^\mu = 0 & \text{radiation gauge} \\ k^\mu a_{\mu\nu} = 0 & \text{Lorentz gauge} \end{cases}$$

Originally we had 10 variables, but the constraints given by the TT gauge reduce this to 2 variables. If we choose to orient the coordinate system such that the wave propagates along the z-axis (in-line with k^μ), the most general shape becomes,

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & b & -a & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} e^{i\omega(z-t)}$$

8.5 Waves

Therefore the direction of propagation is in the **longitudinal** direction with two perpendicular directions of oscillation in the transverse directions.



Appendix

A Connections

Throughout the duration of this course, and in most working theories of gravity, we have taken the connection symbols $\Gamma^{\alpha}{}_{\gamma\beta}$ to be the Christoffel symbols. However, these are a *special case* of the connection symbols with the following two properties,

$$\bar{\Gamma}^{\alpha}{}_{\gamma\beta} = \bar{\Gamma}^{\alpha}{}_{\beta\gamma} \qquad \bar{\nabla}_{\gamma} g_{\alpha\beta} = 0 \tag{8.2}$$

Note that the "bar on top" is used to identify the special case of the Christoffel symbols $\bar{\Gamma}^{\alpha}{}_{\gamma\beta}$ and the Levi-Civita connection $\bar{\nabla}_{\gamma}$ that satisfy these properties. This will be used for this entire section. Albeit throughout the rest of these notes, the bar is omitted as only the Christoffel symbols and the Levi-Civita connection are being considered.

In all scenarios, the general connection acting on a vector w^{α} is *defined* as,

$$\nabla_{\mu} w^{\alpha} \equiv (\partial_{\mu} w^{\alpha} + \Gamma^{\alpha}{}_{\mu\nu} w^{\nu})$$

Where $\Gamma^{\alpha}{}_{\mu\nu}$ are known as the **connection symbols**. How does this work when acting on a tensor?

In order to determine $\nabla_{\mu} C_{\alpha\beta}$ one needs to first come up with an expression for how the general connection acts on a co-vector. Utilizing how it acts on vectors,

$$\nabla_{\mu} w^{\alpha} = (\partial_{\mu} w^{\alpha} + \Gamma^{\alpha}{}_{\mu\nu} w^{\mu}) \tag{8.3}$$

Consider the connection acting on $w^{\alpha} v_{\alpha}$ where v_{β} is a general co-vector. Since $w^{\alpha} v_{\alpha}$ is a contraction of indices, it is just a scalar or in this particular case a scalar function. Therefore the connection acting on this behaves as the normal derivative. Using this and product rule,

$$\nabla_{\mu}(w^{\alpha} v_{\alpha}) = \partial_{\mu}(w^{\alpha} v_{\alpha}) = \partial_{\mu}(w^{\alpha}) v_{\alpha} + w^{\alpha} \partial_{\mu}(v_{\alpha}) \tag{8.4}$$

Alternatively, one can ignore the contraction and apply product rule directly using the connection,

$$\nabla_{\mu}(w^{\alpha} v_{\alpha}) = \nabla_{\mu}(w^{\alpha}) v_{\alpha} + w^{\alpha} \nabla_{\mu}(v_{\alpha})$$

Which by (8.3) gives,

$$\nabla_{\mu}(w^{\alpha} v_{\alpha}) = (\partial_{\mu} w^{\alpha} + \Gamma^{\alpha}{}_{\mu\nu} w^{\mu}) v_{\alpha} + w^{\alpha} \nabla_{\mu}(v_{\alpha}) \tag{8.5}$$

Equating (8.4) with (8.5) gives,

$$\partial_{\mu}(w^{\alpha}) v_{\alpha} + w^{\alpha} \partial_{\mu}(v_{\alpha}) = (\partial_{\mu} w^{\alpha} + \Gamma^{\alpha}{}_{\mu\nu} w^{\mu}) v_{\alpha} + w^{\alpha} \nabla_{\mu}(v_{\alpha})$$

Canceling terms gives,

$$w^{\alpha} \partial_{\mu}(v_{\alpha}) - v_{\alpha} \Gamma^{\alpha}{}_{\mu\nu} w^{\mu} = w^{\alpha} \nabla_{\mu}(v_{\alpha})$$

Relabeling indices,

$$w^{\alpha} \partial_{\mu}(v_{\alpha}) - w^{\alpha} \Gamma^{\gamma}{}_{\alpha\nu} v_{\gamma} = w^{\alpha} \nabla_{\mu}(v_{\alpha})$$

Since this is true for arbitrary w^{α}, it can be factored out; revealing the desired result,

$$\nabla_{\mu} v_{\alpha} = \partial_{\mu} v_{\alpha} - \Gamma^{\gamma}{}_{\alpha\nu} v_{\gamma}$$

Now in order to determine how the covariant derivative acts on a general tensor $C_{\alpha\beta}$, all that needs to be done is generalize the procedure by contracting $C_{\alpha\beta}$ down to scalar function using general vectors and co-vectors $C_{\alpha\beta} w^{\alpha} v^{\beta}$. This gives the relation,

$$\begin{aligned}
\nabla_{\mu}\left(C_{\alpha\beta} w^{\alpha} v^{\beta}\right) &= \nabla_{\mu}(C_{\alpha\beta}) w^{\alpha} v^{\beta} + C_{\alpha\beta} \nabla_{\mu}(w^{\alpha}) v^{\beta} + C_{\alpha\beta} w^{\alpha} \nabla_{\mu}\left(v^{\beta}\right) \\
&= \nabla_{\mu}(C_{\alpha\beta}) w^{\alpha} v^{\beta} + C_{\alpha\beta}(\partial_{\mu} w^{\alpha} + \Gamma^{\alpha}{}_{\mu\nu} w^{\mu}) v^{\beta} + C_{\alpha\beta} w^{\alpha}\left(\partial_{\mu} v^{\beta} + \Gamma^{\beta}{}_{\mu\nu} v^{\mu}\right) \\
&= \partial_{\mu}\left(C_{\alpha\beta} w^{\alpha} v^{\beta}\right)
\end{aligned}$$

$$= \partial_\mu(C_{\alpha\beta})w^\alpha v^\beta + C_{\alpha\beta}\partial_\mu(w^\alpha)v^\beta + C_{\alpha\beta}w^\alpha\partial_\mu(v^\beta)$$

Canceling terms gives,

$$\nabla_\mu(C_{\alpha\beta})w^\alpha v^\beta + C_{\alpha\beta}\Gamma^\alpha{}_{\mu\nu}w^\mu v^\beta + C_{\alpha\beta}w^\alpha\Gamma^\beta{}_{\mu\nu}v^\mu = \partial_\mu(C_{\alpha\beta})w^\alpha v^\beta$$

Relabel indices,

$$\nabla_\mu(C_{\alpha\beta})w^\alpha v^\beta + C_{\gamma\beta}\Gamma^\gamma{}_{\mu\alpha}w^\alpha v^\beta + C_{\alpha\xi}\Gamma^\xi{}_{\mu\beta}w^\alpha v^\beta = \partial_\mu(C_{\alpha\beta})w^\alpha v^\beta$$

Since w^α and v^β are completely unspecified, they can be factored out giving the required result,

$$\nabla_\mu C_{\alpha\beta} = \partial_\mu C_{\alpha\beta} - C_{\gamma\beta}\Gamma^\gamma{}_{\mu\alpha} - C_{\alpha\xi}\Gamma^\xi{}_{\mu\beta} \tag{8.6}$$

Obviously, this procedure can be generalized to,

$$\begin{aligned}\nabla_\mu T^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} &= \partial_\mu T^{\alpha_1\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} + \cdots \\ &\cdots + \Gamma^{\alpha_1}{}_{\mu\sigma}T^{\sigma\alpha_2\cdots\alpha_p}{}_{\beta_1\cdots\beta_q} + \cdots + \Gamma^{\alpha_p}{}_{\mu\sigma}T^{\sigma\alpha_1\cdots\alpha_{p-1}\sigma}{}_{\beta_1\cdots\beta_q} + \cdots \\ &\cdots - \Gamma^\sigma{}_{\mu\beta_1}T^{\alpha_1\cdots\alpha_p}{}_{\sigma\beta_2\cdots\beta_q} - \cdots - \Gamma^\sigma{}_{\mu\beta_q}T^{\alpha_1\cdots\alpha_1}{}_{\beta_1\cdots\beta_{q-1}\sigma}\end{aligned}$$

This is true for the *general connection* as well as the Levi-Civita connection by proxy. We hope that $\nabla_\alpha w^\beta$ itself remains a tensor. Therefore it *must* transform well under a change of coordinates. Enforcing this, we can prove that the connection coefficients Γ are **not** a tensor in general.

Suppose $\nabla_\alpha w^\beta$ was originally expressed in terms of the set of coordinates $\{x^\mu\}$. The new tensor expression in terms of $\{\tilde{x}^\mu\}$ is given by,

$$\widetilde{\nabla_\alpha w^\beta} = \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\nabla_\gamma w^\eta$$

Using (8.3),

$$\widetilde{\partial_\alpha w^\beta} + \widetilde{\Gamma^\beta{}_{\alpha\delta}w^\delta} = \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}(\partial_\gamma w^\eta + \Gamma^\eta{}_{\gamma\sigma}w^\sigma)$$

Examining the $\partial_\gamma w^\eta$ term, since w^η is a tensor it transforms well. Using this idea and product rule then chain rule,

$$\partial_\gamma w^\eta = \partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\beta}\tilde{w}^\beta\right) = \partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\zeta}\right)\tilde{w}^\zeta + \frac{\partial x^\eta}{\partial \tilde{x}^\beta}\partial_\gamma(\tilde{w}^\beta) = \partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\zeta}\right)\tilde{w}^\zeta + \frac{\partial x^\eta}{\partial \tilde{x}^\beta}\frac{\partial \tilde{w}^\beta}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\alpha}{\partial x^\gamma}$$

Subbing this in,

$$\widetilde{\partial_\alpha w^\beta} + \widetilde{\Gamma^\beta{}_{\alpha\delta}w^\delta} = \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\zeta}\right)\tilde{w}^\zeta + \frac{\partial \tilde{w}^\beta}{\partial \tilde{x}^\alpha} + \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\Gamma^\eta{}_{\gamma\sigma}w^\sigma$$

But $\widetilde{\partial_\alpha w^\beta}$ is precisely $\frac{\partial \tilde{w}^\beta}{\partial \tilde{x}^\alpha}$ therefore,

$$\widetilde{\Gamma^\beta{}_{\alpha\delta}w^\delta} = \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\zeta}\right)\tilde{w}^\zeta + \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\Gamma^\eta{}_{\gamma\sigma}w^\sigma$$

Relabel indices,

$$\tilde{\Gamma}^\beta{}_{\alpha\delta}\tilde{w}^\delta = \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\delta}\right)\tilde{w}^\delta + \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\Gamma^\eta{}_{\gamma\sigma}\frac{\partial x^\sigma}{\partial \tilde{x}^\delta}\tilde{w}^\delta$$

Since $\tilde{w}^\delta$ is arbitrary,

$$\tilde{\Gamma}^\beta{}_{\alpha\delta} = \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\delta}\right) + \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\frac{\partial x^\sigma}{\partial \tilde{x}^\delta}\Gamma^\eta{}_{\gamma\sigma}$$

If $\Gamma^\eta{}_{\gamma\sigma}$ were to transform well, it would be required that

$$\tilde{\Gamma}^\beta{}_{\alpha\delta} = \frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\frac{\partial x^\sigma}{\partial \tilde{x}^\delta}\Gamma^\eta{}_{\gamma\sigma}$$

However, the extra term $\frac{\partial x^\gamma}{\partial \tilde{x}^\alpha}\frac{\partial \tilde{x}^\beta}{\partial x^\eta}\partial_\gamma\left(\frac{\partial x^\eta}{\partial \tilde{x}^\delta}\right)$ prevents Γ from transforming well. Therefore Γ is not a tensor.

Now how does parallel transport work under the general connection? Recall that we say a general vector w^α is parallel transported along a geodesic with tangent X iff $\nabla_X w^\alpha = 0$. If the scalar product $w_1^\mu g_{\mu\nu} w_2^\nu$ is to be left invariant under parallel transport, it must be that,

$$0 = \nabla_X(w_1^\mu g_{\mu\nu} w_2^\nu) = X^\beta \nabla_\beta(w_1^\mu g_{\mu\nu} w_2^\nu)$$

Although $w_1^\mu g_{\mu\nu} w_2^\nu$ is in fact a scalar field (due to the contraction) we can and will ignore this fact an apply the product rule anyway,

$$0 = X^\beta \nabla_\beta(w_1^\mu) g_{\mu\nu} w_2^\nu + w_1^\mu X^\beta \nabla_\beta(g_{\mu\nu}) w_2^\nu + w_1^\mu g_{\mu\nu} X^\beta \nabla_\beta(w_2^\nu)$$

But w_1 and w_2 are individually parallel transported along $\mathcal{C}$ which gives $\nabla_\beta(w_1^\mu) = 0$ and $\nabla_\beta(w_2^\nu) = 0$. Thus,

$$0 = w_1^\mu X^\beta \nabla_\beta(g_{\mu\nu}) w_2^\nu$$

Since w_1 and w_2 are any vectors and the curve $\mathcal{C}$ is arbitrary and thus by extension X^β is arbitrary, they may also be factored out,

$$0 = \nabla_\beta g_{\mu\nu} \tag{8.7}$$

Therefore if the scalar product is to be left invariant under parallel transport (as it should), it must be that (8.7). If this holds true, the connection is said to be **metric compatible**. The Levi-Civita symbols are an example of a connection that is metric compatible. When considering the subset of metric compatible connections, one can write the connection symbols in terms of the metric and the **torsion tensor**,

$$T^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\beta\gamma} - \Gamma^\alpha{}_{\gamma\beta}$$

Recall that the Christoffel symbols are symmetric in the lower indices (8.2). As such, the Christoffel symbols are said to be *torsion free* as $\bar{T}^\alpha{}_{\beta\gamma} = 0$. By relabeling the indices of (8.7),

$$\nabla_\mu g_{\beta\gamma} = 0 \quad \nabla_\beta g_{\gamma\mu} = 0 \quad \nabla_\gamma g_{\mu\beta} = 0$$

Which gives three relations using (8.6),

$$\begin{aligned}
0 = \nabla_\mu g_{\beta\gamma} = \nabla_\mu g_{\beta\gamma} = \partial_\mu g_{\beta\gamma} - g_{\sigma\gamma}\Gamma^\sigma{}_{\mu\beta} - g_{\beta\sigma}\Gamma^\sigma{}_{\mu\gamma} \\
0 = \nabla_\beta g_{\gamma\mu} = \nabla_\beta g_{\gamma\mu} = \partial_\beta g_{\gamma\mu} - g_{\sigma\mu}\Gamma^\sigma{}_{\beta\gamma} - g_{\gamma\sigma}\Gamma^\sigma{}_{\beta\mu} \\
0 = \nabla_\gamma g_{\mu\beta} = \nabla_\gamma g_{\mu\beta} = \partial_\gamma g_{\mu\beta} - g_{\sigma\beta}\Gamma^\sigma{}_{\gamma\mu} - g_{\mu\sigma}\Gamma^\sigma{}_{\gamma\beta}
\end{aligned}$$

Adding the first two and subtracting the third gives,

$$0 = \partial_\mu g_{\beta\gamma} - g_{\sigma\gamma}\Gamma^\sigma{}_{\mu\beta} - g_{\beta\sigma}\Gamma^\sigma{}_{\mu\gamma} + \partial_\beta g_{\gamma\mu} - g_{\sigma\mu}\Gamma^\sigma{}_{\beta\gamma} - g_{\gamma\sigma}\Gamma^\sigma{}_{\beta\mu} - \partial_\gamma g_{\mu\beta} + g_{\sigma\beta}\Gamma^\sigma{}_{\gamma\mu} + g_{\mu\sigma}\Gamma^\sigma{}_{\gamma\beta}$$

Using the torsion tensor $T^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\beta\gamma} - \Gamma^\alpha{}_{\gamma\beta}$ and the symmetric symbols $\Gamma^\alpha{}_{(\beta\gamma)} = \frac{1}{2}(\Gamma^\alpha{}_{\beta\gamma} + \Gamma^\alpha{}_{\gamma\beta})$

$$0 = \partial_\mu g_{\beta\gamma} + \partial_\beta g_{\gamma\mu} - \partial_\gamma g_{\mu\beta} - 2g_{\sigma\gamma}\Gamma^\sigma{}_{(\mu\beta)} + g_{\sigma\beta}T^\sigma{}_{\gamma\mu} + g_{\mu\sigma}T^\sigma{}_{\gamma\beta}$$

In order to determine what $\partial_\mu g_{\beta\gamma} + \partial_\beta g_{\gamma\mu} - \partial_\gamma g_{\mu\beta}$ is, look at the special case of the Christoffel symbols $\bar{\Gamma}$. If we replace the general connection Γ with the Christoffel connection symbols, $\bar{\Gamma}$, then the torsion terms drop out because $\bar{\Gamma}$ is torsion-less, or that they are symmetric in the bottom two indices, $\bar{\Gamma}^\alpha{}_{\beta\gamma} = \bar{\Gamma}^\alpha{}_{\gamma\beta}$. Therefore,

$$0 = \partial_\mu g_{\beta\gamma} + \partial_\beta g_{\gamma\mu} - \partial_\gamma g_{\mu\beta} - 2g_{\sigma\gamma}\bar{\Gamma}^\sigma{}_{(\mu\beta)}$$

$$g_{\sigma\gamma}\bar{\Gamma}^\sigma{}_{\mu\beta} = \frac{1}{2}(\partial_\mu g_{\beta\gamma} + \partial_\beta g_{\gamma\mu} - \partial_\gamma g_{\mu\beta})$$

Therefore,

$$0 = 2g_{\sigma\gamma}\bar{\Gamma}^\sigma{}_{\mu\beta} - 2g_{\sigma\gamma}\Gamma^\sigma{}_{(\mu\beta)} + g_{\sigma\beta}T^\sigma{}_{\gamma\mu} + g_{\mu\sigma}T^\sigma{}_{\gamma\beta}$$

$$g_{\sigma\gamma}\Gamma^{\sigma}{}_{(\mu\beta)} = g_{\sigma\gamma}\bar{\Gamma}^{\sigma}{}_{\mu\beta} + \frac{1}{2}(g_{\sigma\beta}T^{\sigma}{}_{\gamma\mu} + g_{\mu\sigma}T^{\sigma}{}_{\gamma\beta})$$

Multiply by inverse metric $g^{\sigma\gamma}$,

$$\Gamma^{\alpha}{}_{(\mu\beta)} = \bar{\Gamma}^{\alpha}{}_{\mu\beta} + \frac{1}{2}(g^{\alpha\gamma}g_{\sigma\beta}T^{\sigma}{}_{\gamma\mu} + g^{\alpha\gamma}g_{\mu\sigma}T^{\sigma}{}_{\gamma\beta})$$

Then lowering and raising corresponding indices,

$$\Gamma^{\alpha}{}_{(\mu\beta)} = \bar{\Gamma}^{\alpha}{}_{\mu\beta} + \frac{1}{2}(g^{\alpha\gamma}T_{\beta\gamma\mu} + g^{\alpha\gamma}T_{\mu\gamma\beta})$$

$$\Gamma^{\alpha}{}_{(\mu\beta)} = \bar{\Gamma}^{\alpha}{}_{\mu\beta} + \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\mu} + T_{\mu}{}^{\alpha}{}_{\beta}\right)$$

Relabeling indices gives the required result,

$$\Gamma^{\alpha}{}_{(\beta\gamma)} = \bar{\Gamma}^{\alpha}{}_{\beta\gamma} + \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\gamma} + T_{\gamma}{}^{\alpha}{}_{\beta}\right) \tag{8.8}$$

The general connection coefficients are related to the Christoffel symbols in this way. This leads to the general shape for an **affine connection**.

Just using the definition, $\Gamma^{\alpha}{}_{(\beta\gamma)} = \frac{1}{2}(\Gamma^{\alpha}{}_{\beta\gamma} + \Gamma^{\alpha}{}_{\gamma\beta})$ (8.8) becomes,

$$\frac{1}{2}(\Gamma^{\alpha}{}_{\beta\gamma} + \Gamma^{\alpha}{}_{\gamma\beta}) = \bar{\Gamma}^{\alpha}{}_{\beta\gamma} + \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\gamma} + T_{\gamma}{}^{\alpha}{}_{\beta}\right)$$

$$\frac{1}{2}\Gamma^{\alpha}{}_{\gamma\beta} = \bar{\Gamma}^{\alpha}{}_{\beta\gamma} + \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\gamma} + T_{\gamma}{}^{\alpha}{}_{\beta} - \Gamma^{\alpha}{}_{\beta\gamma}\right)$$

$$\frac{1}{2}\Gamma^{\alpha}{}_{\gamma\beta} = \bar{\Gamma}^{\alpha}{}_{\beta\gamma} + \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\gamma} + T_{\gamma}{}^{\alpha}{}_{\beta} - \Gamma^{\alpha}{}_{\beta\gamma} + \Gamma^{\alpha}{}_{\gamma\beta} - \Gamma^{\alpha}{}_{\gamma\beta}\right)$$

$$\Gamma^{\alpha}{}_{\gamma\beta} = \bar{\Gamma}^{\alpha}{}_{\beta\gamma} + \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\gamma} + T_{\gamma}{}^{\alpha}{}_{\beta} - \Gamma^{\alpha}{}_{\beta\gamma} + \Gamma^{\alpha}{}_{\gamma\beta}\right)$$

By definition of the torsion tensor and symmetric indices in $\bar{\Gamma}$

$$\Gamma^{\alpha}{}_{\gamma\beta} = \bar{\Gamma}^{\alpha}{}_{\gamma\beta} + \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\gamma} + T_{\gamma}{}^{\alpha}{}_{\beta} + T^{\alpha}{}_{\gamma\beta}\right)$$

Here the extra terms are known as the **contorsion tensor** $K^{\alpha}{}_{\gamma\beta}$,

$$K^{\alpha}{}_{\gamma\beta} = \frac{1}{2}\left(T_{\beta}{}^{\alpha}{}_{\gamma} + T_{\gamma}{}^{\alpha}{}_{\beta} + T^{\alpha}{}_{\gamma\beta}\right)$$

B Lie Derivative

In section (5.5) we introduced the notion of a covariant derivative. There is another derivative of importance: the **Lie Derivative**. It can characterize where tensors are invariant under translational shifts.

When considering a linear transformation from the coordinates of a point x^{μ} to another set of coordinates $\tilde{x}^{\mu}$ given by,

$$\tilde{x}^{\mu} = x^{\mu} + \epsilon\xi^{\mu} \tag{8.9}$$

This will modify the representation or coordinates of fields that depend on those points x^{μ}. For a scalar field $\phi(x^{\mu})$ the change in field under this transformation (8.9) is given by the directional derivative (in the direction ξ) as in multi-variable calculus. At a generic point x^{γ},

$$\Delta_{\xi}\phi(x^{\gamma}) = \xi^{\mu}\partial_{\mu}\phi(x^{\gamma}) \tag{8.10}$$

However on a generic vector *field* W^μ, the change in the representation of W is given by,

$$W^\mu(\tilde{x}^\gamma) - \tilde{W}^\mu(\tilde{x}^\gamma) \tag{8.11}$$

Essentially, this tells one how the new vector field compares to the old vector field at a particular point $\tilde{x}^\gamma$. The vector field is left invariant if this quantity becomes zero. In order to determine this quantity explicitly, we can utilize how the vector field changes under the translation (8.9), ignoring that the transformation can be viewed as affecting the actual components of the vector field. Using a first order approximation for this Taylor expansion in small ϵ,

$$W^\mu(\tilde{x}^\gamma) = W^\mu(x^\gamma) + \epsilon\xi^\nu\partial_\nu W^\mu(x^\gamma) + O^\mu\left(\epsilon^2\right) \tag{8.12}$$

Notice that this expansion has no interest in the transformed vector field $\tilde{W}^\mu$. I am taking $O^\mu\left(\epsilon^2\right) \approx 0$ from now on. In order to determine how the vector field itself is modified, use the fact that W^μ is a tensor field that transforms well using a Jacobian,

$$\tilde{W}^\mu(\tilde{x}^\gamma) = \frac{\partial\tilde{x}^\mu}{\partial x^\nu}W^\nu(x^\gamma)$$

The Jacobian $\frac{\partial\tilde{x}^\mu}{\partial x^\nu}$ can be computed by (8.9) directly,

$$\frac{\partial\tilde{x}^\mu}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}(x^\mu + \epsilon\xi^\mu) = \frac{\partial}{\partial x^\nu}(x^\mu) + \epsilon\frac{\partial}{\partial x^\nu}(\xi^\mu) = \delta^\mu{}_\nu + \epsilon\partial_\nu\xi^\mu$$

Therefore,

$$\tilde{W}^\mu(\tilde{x}^\gamma) = [\delta^\mu{}_\nu + \epsilon\partial_\nu\xi^\mu]W^\nu(x^\gamma) = W^\mu(x^\gamma) + \epsilon W^\nu\partial_\nu\xi^\mu(x^\gamma) \tag{8.13}$$

Therefore (8.11) can be obtained by taking (8.12) subtract (8.13),

$$W^\mu(\tilde{x}^\gamma) - \tilde{W}^\mu(\tilde{x}^\gamma) = \epsilon\xi^\nu\partial_\nu W^\mu(x^\gamma) - \epsilon W^\nu\partial_\nu\xi^\mu(x^\gamma)$$

Taking this difference per unit length ϵ under the limit for $\epsilon \to 0$ gives the definition of the *Lie derivative* in the direction of ξ^μ.

$$L_\xi W^\mu(x^\gamma) = \xi^\nu\partial_\nu W^\mu(x^\gamma) - W^\nu\partial_\nu\xi^\mu(x^\gamma)$$

Or without writing dependence on x^γ,

$$L_\xi W^\mu = \xi^\nu\partial_\nu W^\mu - W^\nu\partial_\nu\xi^\mu \tag{8.14}$$

Symmetries in W^μ under (8.9) are known as isometries and are exposed by $L_\xi W^\mu = 0$. How does this Lie derivative act on the metric tensor? This can be determined by exploiting how the Lie derivative acts on a scalar field ϕ given by (8.10).

$$L_\xi\phi = \xi^\nu\partial_\nu\phi \tag{8.15}$$

Considering two generic vector fields W^μ and V^μ, one can construct a scalar field using the metric $g_{\alpha\beta}$; $\phi = W^\alpha g_{\alpha\beta}V^\beta$. Then by (8.15),

$$L_\xi\phi = L_\xi\left(W^\alpha g_{\alpha\beta}V^\beta\right) = \xi^\gamma\partial_\gamma\left(W^\alpha g_{\alpha\beta}V^\beta\right)$$

By product rule,

$$L_\xi\left(W^\alpha g_{\alpha\beta}V^\beta\right) = \xi^\gamma\partial_\gamma(W^\alpha)g_{\alpha\beta}V^\beta + W^\alpha\xi^\gamma\partial_\gamma(g_{\alpha\beta})V^\beta + W^\alpha g_{\alpha\beta}\xi^\gamma\partial_\gamma\left(V^\beta\right) \tag{8.16}$$

Shifting focus, we should be able to treat $W^\alpha g_{\alpha\beta}V^\beta$ as a tensor field $W^\mu g_{\alpha\beta}V^\nu$ and apply product rule directly,

$$L_\xi\left(W^\alpha g_{\alpha\beta}V^\beta\right) = L_\xi(W^\alpha)g_{\alpha\beta}V^\beta + W^\alpha L_\xi(g_{\alpha\beta})V^\beta + W^\alpha g_{\alpha\beta}L_\xi\left(V^\beta\right)$$

The action of L_ξ on vector fields is known giving,

$$L_\xi\left(W^\alpha g_{\alpha\beta}V^\beta\right) = (\xi^\gamma\partial_\gamma W^\alpha - W^\gamma\partial_\gamma\xi^\alpha)g_{\alpha\beta}V^\beta + W^\alpha L_\xi(g_{\alpha\beta})V^\beta + W^\alpha g_{\alpha\beta}\left(\xi^\gamma\partial_\gamma V^\beta - V^\gamma\partial_\gamma\xi^\beta\right) \tag{8.17}$$

Equating (8.16) with (8.17) and canceling duplicate terms yields,

$$W^\alpha \xi^\gamma \partial_\gamma (g_{\alpha\beta}) V^\beta = -W^\gamma \partial_\gamma \xi^\alpha g_{\alpha\beta} V^\beta + W^\alpha L_\xi (g_{\alpha\beta}) V^\beta - W^\alpha g_{\alpha\beta} V^\gamma \partial_\gamma \xi^\beta$$
$$W^\alpha L_\xi (g_{\alpha\beta}) V^\beta = W^\alpha \xi^\gamma \partial_\gamma (g_{\alpha\beta}) V^\beta + W^\gamma \partial_\gamma \xi^\alpha g_{\alpha\beta} V^\beta + W^\alpha g_{\alpha\beta} V^\gamma \partial_\gamma \xi^\beta$$

Relabel indices,

$$W^\alpha L_\xi (g_{\alpha\beta}) V^\beta = W^\alpha \xi^\gamma \partial_\gamma (g_{\alpha\beta}) V^\beta + W^\alpha \partial_\alpha \xi^\gamma g_{\gamma\beta} V^\beta + W^\alpha g_{\alpha\gamma} V^\beta \partial_\beta \xi^\gamma$$

Since W^α and V^β are arbitrary, they can be factored out,

$$L_\xi (g_{\alpha\beta}) = \xi^\gamma \partial_\gamma (g_{\alpha\beta}) + \partial_\alpha \xi^\gamma g_{\gamma\beta} + g_{\alpha\gamma} \partial_\beta \xi^\gamma$$

Using the comma notation,

$$L_\xi (g_{\alpha\beta}) = g_{\alpha\beta,\gamma} \xi^\gamma + g_{\gamma\beta} \xi^\gamma{}_{,\alpha} + g_{\alpha\gamma} \xi^\gamma{}_{,\beta}$$

As discussed above, isometries are exposed whenever $L_\xi (g_{\alpha\beta}) = 0$, therefore in order for (8.9) to be an isometry, $g_{\alpha\beta}$ needs to satisfy,

$$0 = g_{\alpha\beta,\gamma} \xi^\gamma + g_{\gamma\beta} \xi^\gamma{}_{,\alpha} + g_{\alpha\gamma} \xi^\gamma{}_{,\beta} \tag{8.18}$$

This is a form of the **Killing equation**.

C Killing Vectors

At the end of section B, we obtained an expression for the killing equation (8.18). This section aims to manipulate this result and arrive at the idea of killing vectors ξ^μ and their use in general relativity.

The definition of the Christoffel symbols are,

$$\Gamma^\alpha{}_{\beta\gamma} = \frac{1}{2} g^{\alpha\mu} (g_{\beta\mu,\gamma} + g_{\gamma\mu,\beta} - g_{\gamma\beta,\mu})$$

Which gives,

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

Multiply by ξ^μ

$$2\Gamma^\alpha{}_{\beta\gamma} g_{\alpha\mu} \xi^\mu = g_{\beta\mu,\gamma} \xi^\mu + g_{\gamma\mu,\beta} \xi^\mu - g_{\gamma\beta,\mu} \xi^\mu$$

Sub in (8.18),

$$2\Gamma^\alpha{}_{\beta\gamma} \xi_\alpha = g_{\beta\mu,\gamma} \xi^\mu + g_{\gamma\mu,\beta} \xi^\mu + g_{\mu\beta} \xi^\mu{}_{,\gamma} + g_{\gamma\mu} \xi^\mu{}_{,\beta}$$
$$2\Gamma^\alpha{}_{\beta\gamma} \xi_\alpha = \partial_\gamma (g_{\beta\mu}) \xi^\mu + \partial_\beta (g_{\gamma\mu}) \xi^\mu + g_{\mu\beta} \partial_\gamma (\xi^\mu) + g_{\gamma\mu} \partial_\beta (\xi^\mu)$$

Inverse product rule,

$$2\Gamma^\alpha{}_{\beta\gamma} \xi_\alpha = \partial_\gamma (g_{\beta\mu} \xi^\mu) + \partial_\beta (g_{\gamma\mu} \xi^\mu)$$

Lowering of indices,

$$2\Gamma^\alpha{}_{\beta\gamma} \xi_\alpha = \partial_\gamma (\xi_\beta) + \partial_\beta (\xi_\gamma)$$
$$2\Gamma^\alpha{}_{\beta\gamma} \xi_\alpha = \xi_{\beta,\gamma} + \xi_{\gamma,\beta}$$
$$0 = \xi_{\beta,\gamma} + \xi_{\gamma,\beta} - 2\Gamma^\alpha{}_{\beta\gamma} \xi_\alpha \tag{8.19}$$

This is another form of the killing equation. To take this further, The Levi-Civita connection acting on co-vectors is given by,

$$\nabla_\alpha (\xi_\beta) = \partial_\alpha (\xi_\beta) - \Gamma^\gamma{}_{\alpha\beta} \xi_\gamma = \xi_{\beta,\alpha} - \Gamma^\gamma{}_{\alpha\beta} \xi_\gamma$$

Therefore from (8.19),

$$0 = \xi_{\beta,\gamma} + \xi_{\gamma,\beta} - \Gamma^\alpha{}_{\beta\gamma} \xi_\alpha - \Gamma^\alpha{}_{\beta\gamma} \xi_\alpha$$

Since the Christoffel symbols are symmetric in the lower indices,

$$0 = \xi_{\beta,\gamma} - \Gamma^\alpha{}_{\gamma\beta} \xi_\alpha + \xi_{\gamma,\beta} - \Gamma^\alpha{}_{\beta\gamma} \xi_\alpha$$

$$0 = \nabla_\gamma \xi_\beta + \nabla_\beta \xi_\gamma$$

Or by relabeling indices,

$$0 = \nabla_\alpha \xi_\beta + \nabla_\beta \xi_\alpha \tag{8.20}$$

Which gives the killing equation in a popular form.

But what are killing vectors ξ^α used for? Suppose we have a geodesic with tangent vector V^α and killing vector ξ^β, by product rule of the connection,

$$\nabla_V\left(\xi^\beta V_\beta\right) = \nabla_V\left(\xi_\beta V^\beta\right) = \nabla_V(\xi_\beta)V^\beta + \xi_\beta \nabla_V(V_\beta) = V^\gamma \nabla_\gamma(\xi_\beta)V^\beta + \xi_\beta V^\gamma \nabla_\gamma\left(V^\beta\right)$$

Since V^α is tangent to a geodesic, it must satisfy the geodesic equation,

$$\nabla_V V^\beta = 0$$

Therefore,

$$\nabla_V\left(\xi^\beta V_\beta\right) = V^\gamma \nabla_\gamma(\xi_\beta)V^\beta \tag{8.21}$$

Notice the switch for convenience,

$$\xi^\beta V_\beta = g_{\beta\alpha} g^{\beta\alpha} \xi^\beta V_\beta = \xi_\alpha V^\alpha = \xi_\beta V^\beta$$

But since ξ^β is a killing vector it satisfies (8.20),

$$\nabla_V\left(\xi^\beta V_\beta\right) = -V^\gamma \nabla_\beta(\xi_\gamma)V^\beta$$

By relabeling indices,

$$\nabla_V\left(\xi^\beta V_\beta\right) = -V^\beta \nabla_\gamma(\xi_\beta)V^\gamma$$

By commutativity,

$$\nabla_V\left(\xi^\beta V_\beta\right) = -V^\gamma \nabla_\gamma(\xi_\beta)V^\beta$$

Which by (8.21) gives,

$$\nabla_V\left(\xi^\beta V_\beta\right) = -\nabla_V\left(\xi^\beta V_\beta\right)$$

Which yields,

$$\nabla_V\left(\xi^\beta V_\beta\right) = 0$$

Physically, this means that $\xi^\beta V_\beta$ is a scalar quantity that is **conserved** along the length of the geodesic. As such, one can *construct* killing vectors ξ^α that satisfy (8.20) in order to expose certain components of V^β that are conserved allow the length of the geodesic. As an example, taking $\xi^\alpha = (1,0,0,0)$ as a killing vector, $\xi = \xi^\alpha \partial_\alpha = \partial_0$. If ξ^α is to satisfy the killing equation,

$$0 = \nabla_\alpha \xi_\beta + \nabla_\beta \xi_\alpha$$

Alternatively, use the equivalent form of the killing equation, (8.18),

$$0 = g_{\alpha\beta,\gamma}\xi^\gamma + g_{\gamma\beta}\xi^\gamma{}_{,\alpha} + g_{\alpha\gamma}\xi^\gamma{}_{,\beta}$$

Which for this particular ξ^α, we have,

$$\xi^\gamma{}_{,\alpha} = 0$$

Since ξ^α is only made up of constants. Thus,

$$0 = g_{\alpha\beta,\gamma}\xi^\gamma$$

Which when contracted gives,

$$0 = g_{\alpha\beta,0}\xi^0$$
$$0 = \partial_0 g_{\alpha\beta}$$

Or that $g_{\alpha\beta}$ is time independent,

$$0 = \partial_t g_{\alpha\beta}$$

This concludes a proof that for **stationary metrics**, where $g_{\alpha\beta}$ does not depend on time, the vector $\xi^\alpha = (1,0,0,0)$ is always a killing vector. As such, the time components of V^α are always conserved along the length of a geodesic. This was to be expected because the time component of V^α in the frame of the observed that is moving along the geodesic is constant.