MAU34301 - Differential Geometry
Brief Notes

Liam Kavanagh
B.A. Theoretical Physics

"Differential Geometry is the study of things that are invariant under change of notation."

Mathematical Proverb

1 Manifolds*

1.1 Topological Spaces

A **topological space** is a set X of which the open sets of the topological space satisfy

(a) For any two open sets $X_1, X_2 \subset X$, $X_1 \cap X_2$ and $X_1 \cup X_2$ are also open sets.

(b) The empty set $\emptyset$ and the whole set X are open.

The complement X_1' of any open set X_1 is called a **closed set** of the topological space.

$\mathbb{R}^n$ induces an **induced topology** on a subset $A \subset \mathbb{R}^n$ through the intersections $A \cap U$ where U ranges over all open sets of $\mathbb{R}^n$.

A map $f : X \to Y$ between topological spaces X and Y is **continuous** if the complete inverse image $f^{-1}(U)$ of every open set $U \subseteq Y$ is open in X.

Two topological spaces X and Y are **topologically equivalent**, or **homeomorphic** if there is a bijective map between them such that both it and its inverse are continuous.

The topology on a manifold M:
In every chart U_q, the open regions are open in the topology on M. The totality of open sets on M is found by ensuring all arbitrary unions of collections of these regions is also open. Any open subset $V \subset M$ inherits - i.e. has induced on it - the structure of a manifold: $V = \cup_q V_q$ where $V_q = V \cap U_q$.

A **metric space** is a set equipped with a real-valued "distance function", or metric, $\rho(x, y)$ defined on pairs x, y of elements of the set. The metric satisfies

(a) Symmetry: $\rho(x, y) = \rho(y, x)$,

(b) Positivity: $\rho(x, y) \geq 0$ with equality only for $x = y$,

(c) The triangle inequality: $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$.

*For the purposes of covering the module's examinable material, I am discussing manifolds as they are in the lecture notes. If you would like a more intuitive discussion of the geometry of manifolds, consult Chapter 2 of *Lecture Notes on General Relativity* [1] by Sean M. Carroll.

Example: n-dimensional Euclidean space $\mathbb{R}^n$ is a metric space with metric

$$\rho(x,y) = \sqrt{\sum_{\alpha=1}^{n}(x^\alpha - y^\alpha)^2}. \tag{1.1}$$

A metric space can be induced with a topology by taking as its open sets the unions of open balls. The elements x of an open ball with centre x_0 and radius ε satisfy

$$\rho(x,x_0) < \varepsilon. \tag{1.2}$$

A topological space X is **compact** if every countable collection of open sets covering X contains a finite subcollection also covering X.

A topological space is **connected** if any two points can be joined by a continuous path.

A topological space is called **Hausdorff**† if any two points are contained in disjoint open sets.
All metric spaces are Hausdorff since two open balls with centres x_0 and y_0 and radii $\rho(x_0,y_0)/3$ do not intersect. The manifolds we consider will be Hausdorff spaces.‡

1.2 Manifold Definitions

A **region** is a set $D \subset \mathbb{R}^n$ s.t.

$$\forall P_0 = (x_0^1, ..., x_0^n) \in D, \exists \varepsilon > 0 \text{ s.t. } \left\{P = (x^1,...,x^n) : |x^i - x_0^i| < \varepsilon, i = 1,...,n\right\} \subset D. \tag{1.3}$$

A **differentiable n-dimensional manifold** is a set M which is the union of a collection of subsets U_q with the properties:

(i) Each subset U_q has defined on it coordinates $x_q^\alpha, \alpha = 1,...,n$ called local coordinates. Through these coordinates, U_q looks like a region of Euclidean space with coordinates x_q^α. The subsets U_q together with their local coordinates are called **charts**.

(ii) Each non-empty intersection $U_p \cap U_q$ of charts has defined on it two coordinate systems; (x_p^α) and (x_q^α). We require that the intersection $U_p \cap U_q$ can be expressed in terms of either coordinate system and that they are related in a smooth, one-to-one way. The **transition functions** from x_q^α to x_p^α and vice versa are

$$\begin{aligned} x_p^\alpha &= x_p^\alpha(x_q^1,...,x_q^n), \\ x_q^\alpha &= x_q^\alpha(x_p^1,...,x_p^n). \end{aligned} \tag{1.4}$$

The Jacobian $J_{pq} = \det\left(\dfrac{\partial x_p^\alpha}{\partial x_q^\beta}\right)$ is non-zero on $U_p \cap U_q$.

In Fig. (1), we have $M = \cup_q U_q$. For each chart U_q, there exists a smooth one-to-one map $\phi_q : U_q \to \mathbb{R}^n$, $n = \dim M$, such that $x_q^\alpha = \phi_q(P)^\alpha$ for all $P \in U_q$. Thus, $P = \phi_q^{-1}(x_q)$.

For the intersections $U_p \cap U_q$, we have the relations

$$\begin{aligned} x_q^\alpha &= \phi_q(\phi_p^{-1}(x_p))^\alpha = \left[(\phi_q \circ \phi_p^{-1}(x_p)\right]^\alpha, \\ x_p^\alpha &= \phi_p(\phi_q^{-1}(x_q))^\alpha = \left[(\phi_p \circ \phi_q^{-1}(x_q)\right]^\alpha. \end{aligned} \tag{1.5}$$

†The Hausdorff condition is that any two points can be "housed off" from each other by open sets (pun).
‡Felix Hausdorff lectured at the University of Bonn and in their Mathematics Institute, there is a room named the "Hausdorff-Raum", which is a pun because in German "Raum" can mean both "room" and "space".

An indexed collection $\{(U_q, \phi_q)\}$ of charts U_q along with the maps ϕ_q form an **atlas** of the manifold.

A manifold M is **oriented** if the atlas can be chosen so that for every pair U_p, U_q of intersecting charts the Jacobian J of the transition functions is positive. e.g. $\mathbb{R}^n$ and $\mathcal{S}^n$ are oriented. We say that two coordinate systems x and y define the same orientation of $\mathbb{R}^n$ if $J > 0$ and opposite orientations if $J < 0$.

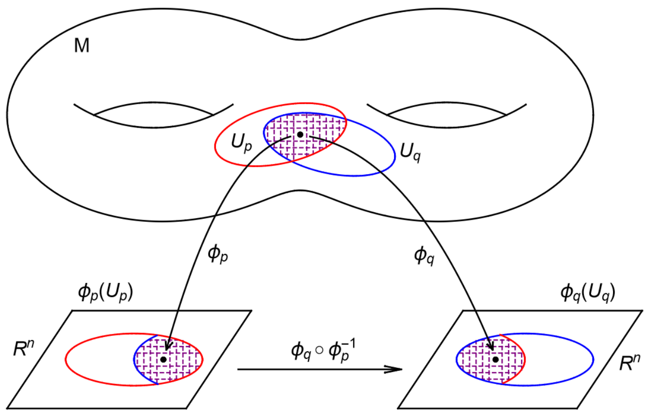

Figure 1: The maps from M to $\mathbb{R}^n$.

1.3 Examples of Manifolds

1. Trivially, any Euclidean space of regions is a manifold.
2. A region of complex space $\mathbb{C}^n$ can be regarded as a region of $\mathbb{R}^{2n}$, so $\mathbb{C}^n$ is a manifold.
3. The circle $\mathcal{S}^1$ is a manifold:

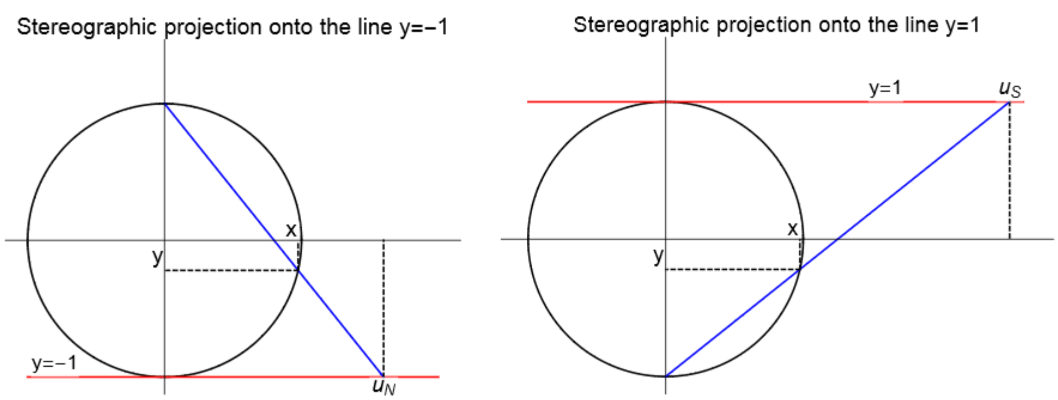

Figure 2: The unit circle embedded in $\mathbb{R}^2$.

Consider Fig. (2), the unit circle embedded in $\mathbb{R}^2$: $x^2 + y^2 = 1$. We introduce two subsets $U_N = \mathcal{S}^1 - (0, 1)$ and $U_S = \mathcal{S}^1 - (0, -1)$. The local coordinates u_N and u_S are obtained using a

stereographic projection onto $y=-1$ and $y=+1$:

$$\begin{aligned} P=(x,y)\to u_N &= \frac{2x}{1-y}, \\ P=(x,y)\to u_S &= \frac{2x}{1+y}. \end{aligned} \tag{1.6}$$

The intersection $U_N \cap U_S$ excludes $x=0$, i.e. $u_N = u_S = 0$.
The transition functions are found from $u_N u_S = \dfrac{2x}{1-y}\dfrac{2x}{1+y} = \dfrac{4x^2}{1-y^2} = \dfrac{4x^2}{x^2} = 4$, and on $U_N \cap U_S$ where $u_N, u_S \neq 0$, we have

$$u_N = \frac{4}{u_S} \implies \frac{\partial u_N}{\partial u_S} = -\frac{4}{u_S^2} \implies J_{NS} = \det\left(\frac{\partial u_N}{\partial u_S}\right) = -\frac{4}{u_S^2} \neq 0. \tag{1.7}$$

4. See `Slides_Diff_Geom-p1.pdf`, p. 21 for the example of $\mathcal{S}^2$.

5. In general, an n-sphere $\mathcal{S}^n$ is a manifold.

6. Given two manifolds $M = \cup_q U_q$, $N = \cup_p V_p$, we can construct their **direct product**

$$M \times N = \cup_{q,p} U_q \times V_p \tag{1.8}$$

where the coordinates on $U_q \times V_p$ are (x_q^α, y_p^β).

1.4 Mappings Between Manifolds

Let $M = \cup_p U_p$ and $N = \cup_q, U_q$ be m and n dimensional manifolds with coordinates (x_p^α) and (y_q^β), respectively.
Suppose a mapping $f: M \to N$ determines functions $y_q^\beta(x_p^1, ..., x_p^m) = f(x_p^1, ..., x_p^m)_q^\beta$. The function f is said to be **smooth** of smoothness class k if for all p, q the functions y_q^β are smooth of smoothness class k, that is, all the partial derivatives up to k^{th} order exist and are continuous. The smoothness class of f cannot exceed the smoothness class of the manifolds.

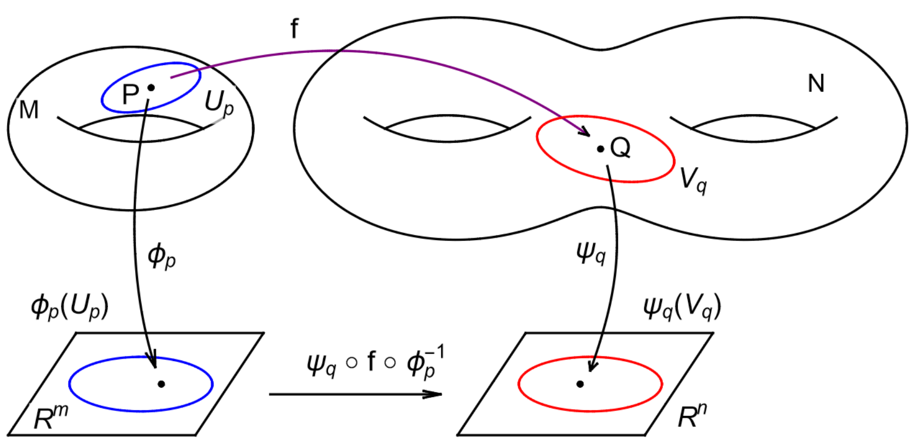

Figure 3:

In Fig. (3), the four maps f, ϕ_p, ψ_q, and $\psi_q \circ f \circ \phi_p^{-1}$ are smooth. We use the notation

$$y_q^\beta = \left[(\psi_q \circ f \circ \phi_p^{-1})(x_p^1, ..., x_p^m)\right]^\beta := f^\beta(x_p^1, ..., x_p^m) = y_q^\beta(x_p^1, ..., x_p^m). \tag{1.9}$$

Notice that since we have maps from the manifolds to Euclidean space, we can use analysis on them. We can construct the map $f: M \to N$ from the charts as $(\psi_q \circ f \circ \phi_p^{-1}) : \mathbb{R}^m \to \mathbb{R}^n$. Since this is a map between Euclidean spaces, we can differentiate normally. This is a notational shortcut which is really just

$$\frac{\partial f}{\partial x^\alpha} = \frac{\partial}{\partial x^\alpha}(\psi_q \circ f \circ \phi_p^{-1})(x_p^1, ..., x_p^m). \tag{1.10}$$

If $N = \mathbb{R}$ is the real line, then f is a real-valued function of the points of M.

The manifolds M and N are said to be **smoothly equivalent** or **diffeomorphic** if there exists a bijective map f such that both $f: M \to N$ and $f^{-1}: N \to M$ exist and are smooth of class $k \geq 1$.

Since f^{-1} exists, the Jacobian $J_{pq} = \det\left(\frac{\partial y_q^\beta}{\partial x_p^\alpha}\right)$ is non-zero everywhere.

1.5 Tangent Spaces

Let $x = x(\tau)$, $a \leq \tau \leq b$ be a curve segment on a manifold $M \ni x(\tau)$. In any chart U_p with coordinates x_p^α, the curve is described by the parametric equations

$$x_p^\alpha = x_p^\alpha(\tau), \ \alpha = 1, ..., m. \tag{1.11}$$

The **velocity vector** tangent to the curve is

$$v := \dot{x} = (\dot{x}_p^1, ..., \dot{x}_p^m). \tag{1.12}$$

In the intersection $U_p \cap U_q$, we have two equally good coordinate systems $x_p^\alpha(\tau)$ and $x_q^\beta(\tau)$ such that

$$x_p^\alpha(x_q^1(\tau), ..., x_q^m(\tau)) = x_p^\alpha(\tau). \tag{1.13}$$

The velocities are thus related by (sum over β)

$$\dot{x}_p^\alpha = \frac{\partial x_p^\alpha}{\partial x_q^\beta}\dot{x}_q^\beta. \tag{1.14}$$

A **tangent vector** to a m-dim manifold M at some point x is represented in terms of local coordinates x_p^α by the m-tuple (ξ^α) of components of the vector such that in any other local coordinate system

$$\xi_p^\alpha = \frac{\partial x_p^\alpha}{\partial x_q^\beta}\xi_q^\beta. \tag{1.15}$$

The set of all tangent vectors to a manifold M at x forms a vector space of dimension $m = \dim M$, called the **tangent space** to M at x, denoted $T_x = T_x M$.

The velocity vector at x of any smooth curve of M through x is a tangent vector to M at x.

From (1.15), we see that the operators $\frac{\partial}{\partial x^\alpha}$ form a basis $e_\alpha = \frac{\partial}{\partial x^\alpha}$ of one-forms for the tangent space T_x:

$$\xi_p^\alpha \frac{\partial}{\partial x_p^\alpha} = \xi_q^\beta \frac{\partial}{\partial x_q^\beta} \tag{1.16}$$

As with the manifolds themselves, we can form maps between tangent spaces of the manifolds.
A smooth map $f: M \to N$ gives rise to a **push-forward** or **induced linear** map of tangent spaces

$$f_*: T_x M \to T_{f(x)} N \tag{1.17}$$

defined as sending the velocity vector at x of some smooth curve $x(\tau)$ on M to the velocity vector at $f(x)$ of the curve $f(x(\tau))$ on N. In terms of local coordinates x^α on M and y^β on N, the push-forward map is

$$\xi^\alpha \to \eta^\beta = \frac{\partial f^\beta}{\partial x^\alpha}\xi^\alpha. \tag{1.18}$$

For a real-valued function $f : M \to \mathbb{R}$, the push-forward map f_* is a real-valued linear function

$$\xi^\alpha \to \eta = \frac{\partial f}{\partial x^\alpha}\xi^\alpha. \tag{1.19}$$

which is the gradient of f at x; a co-vector or one-form. So, f_* an be identified with the differential one-form df:

$$dx_p^\alpha : \xi^\alpha \to \eta = \xi_p^\alpha. \tag{1.20}$$

A **Riemann metric** on a manifold M is a point-dependent, positive-definite quadratic form on the tangent vectors at each point, depending smoothly on the local coordinates. A **pseudo-Riemann metric** requires the quadratic form to be non-degenerate (non-zero determinant).
At each point $x = (x_p^1, ..., x_p^m)$ of each chart U_p, the metric is given by a symmetric matrix $(g_{\alpha\beta}^{(p)}(x_p^1, ..., x_p^m))$ and determines a symmetric scalar product of pairs of tangent vectors at x:

$$\langle \xi, \eta \rangle = g_{\alpha\beta}^{(p)} \xi_p^\alpha \eta_p^\beta = \langle \eta, \xi \rangle, \ \langle \xi, \xi \rangle = |\xi|^2. \tag{1.21}$$

Since the scalar product is a scalar, it is coordinate independent. The coefficients $g_{\alpha\beta}^{(p)}$ transform as

$$g_{\gamma\delta}^{(q)} = \frac{\partial x_p^\alpha}{\partial x_q^\gamma}\frac{\partial x_p^\beta}{\partial x_q^\delta} g_{\alpha\beta}^{(p)}. \tag{1.22}$$

We can rewrite this in terms of basis one-forms in a coordinate-independent way as

$$ds^2 = g_{\alpha\beta}^{(p)} dx_p^\alpha dx_p^\beta = g_{\alpha\beta}^{(q)} dx_q^\alpha dx_q^\beta. \tag{1.23}$$

ds is the **line element** and is used to measure distances between two infinitesimally close points.

A **tensor of type** (k, l) and **rank** $k + l$ on a manifold M of $\dim m$ is given in a local coordinate system (x_p^i) by a family of functions ${}^{(p)}T^{i_1 \dots i_k}_{j_1 \dots j_l}(x)$. In another coordinate system, the components are

$${}^{(q)}T^{s_1 \dots s_k}_{t_1 \dots t_l} = \frac{\partial x_q^{s_1}}{\partial x_p^{i_1}} \cdots \frac{\partial x_q^{s_k}}{\partial x_p^{i_k}} \cdot \frac{\partial x_p^{j_1}}{\partial x_q^{t_1}} \cdots \frac{\partial x_p^{j_l}}{\partial x_q^{t_l}} . {}^{(p)}T^{i_1 \dots i_k}_{j_1 \dots j_l}. \tag{1.24}$$

This is the transformation rule for tensors of type (k, l). See `Slides_Diff_Geom-p1.pdf`, pp. 56-66 for examples.

2 Lie Groups and Lie Algebras

2.1 Lie Groups

A manifold G is called a **Lie group** if it has given on it a group operation with the property that the maps

$$\begin{aligned} \varphi : G \to G, \quad \varphi(g) = g^{-1} \text{ (the inverse)}, \\ \psi : G \times G \to G, \quad \psi(g,h) = gh \text{ (group multiplication)} \end{aligned} \tag{2.1}$$

are smooth maps.

Let G be a Lie group, with the point $g_0 = 1$ the identity element of G and $T = T_{(1)}$ the tangent space at the identity. We can express the group operations on G in a chart U_0 containing g_0 in terms of local coordinates.

We choose coordinates in U_0 so that $g_0 = (0, ..., 0)$ is the origin. We let

$$g_1 = (x^1, ..., x^n), \quad g_2 = (y^1, ..., y^n), \quad g_3 = (z^1, ..., z^n) \tag{2.2}$$

such that $g_k^\sigma g_{k'}^{\sigma'} \in U_0$ for all combinations $k, k' = 1, 2, 3$ and $\sigma, \sigma' = -1, 1$, where $+1$ is the element itself, and -1 is the inverse of the element. Then we have the coordinates

$$g_1 g_2 = \left(\psi^1(x,y), ..., \psi^n(x,y)\right) = \left(\psi^i(x,y)\right), \;\; \psi^i(x,y) = \psi^i(x^1, ..., x^n, y^1, ..., y^n) \tag{2.3}$$

$$g_1^{-1} = \left(\varphi^1(x), ..., \varphi^n(x)\right) = \left(\varphi^i(x)\right), \;\; \varphi^i(x) = \varphi^i(x^1, ..., x^n). \tag{2.4}$$

Here, $\psi(x,y)$ and $\varphi(x)$ satisfy, for all $i = 1, ..., n$:

1. $\psi^i(x,0) = \psi^i(0,x) = x^i$ (multiplication by the identity),
2. $\psi^i(x, \varphi(x)) = 0$ (multiplication of an element and its inverse is the identity),
3. $\psi^i(x, \psi(y,z)) = \psi^i(\psi(x,y), z)$ (associativity).

We can find an expression for $\psi(x,y)$ in a neighbourhood around the identity by Taylor expanding. We find that

$$\psi^i(x,y) = x^i + y^i + b^i_{jk} x^j y^k + \mathcal{O}\left(\{x,y,z\}^3\right), \;\; b^i_{jk} = \left.\frac{\partial^2 \psi^i}{\partial x^j \partial y^k}\right|_{x=y=0}. \tag{2.5}$$

Let $\xi, \eta \in T$, the tangent space of the identity, and let their components in terms of x^i be ξ^i, η^i. The **commutator** $[\xi, \eta] \in T$ is defined by

$$[\xi, \eta]^i = c^i_{jk} \xi^j \eta^k, \quad c^i_{jk} := b^i_{jk} - b^i_{kj}. \tag{2.6}$$

It has the properties

1. $[,]$ is a bilinear operation on the n-dim vector space T,
2. it is skew-symmetric: $[\xi, \eta] = -[\eta, \xi]$,
3. it satisfies Jacobi's Identity: $[[\xi, \eta], \zeta] + [[\zeta, \xi], \eta] + [[\eta, \zeta], \xi] = 0$.

The proof is long and arduous algebra, and can be found in `Slides_Diff_Geom-p3.pdf`, pp. 8-11.

A **Lie algebra** is a vector space $\mathcal{G}$ over a field F with a bilinear operation $[\cdot, \cdot] : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ called a commutator, or Lie bracket. It satisfies the following axioms:

1. $\forall x, y \in \mathcal{G}, [x,x] = 0$ and $[x,y] = -[y,x]$ (skew-symmetry),

2. $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$ (Jacobi identity).

Thus, the tangent space $T=T_{g_0}$ at the identity of the Lie group G, along with the commutator operation, is a Lie algebra called the **Lie algebra of the Lie group** G.

Let $e_i=\dfrac{\partial}{\partial x^i}, i=1,...n$ be the standard basis vectors of T. Then, we have

$$[\xi,\eta]=[\xi,\eta]^i e_i=c^i_{jk}\xi^j\eta^k e_i. \tag{2.7}$$

Choosing $\xi=e_j,\eta=e_k$ and noting that $(e_m)^n=\delta^n_m$, we get

$$[e_j,e_k]=c^i_{jk}e_i. \tag{2.8}$$

The skew-symmetric constants $c^i_{jk}=-c^i_{kj}$ which determine the commutation operation of the Lie algebra are called the **structure constants** of the Lie algebra.

2.2 One-Parameter Subgroups

A **one-parameter subgroup** of a Lie group G is a curve $F(t)$, parametrised by $t\in\mathbb{R}$, on the manifold G such that

$$F(0)=1,\quad F(t_1+t_2)=F(t_1)F(t_2),\quad F(-t)=F(t)^{-1} \tag{2.9}$$

The velocity vector at $F(t)$ is

$$\frac{dF}{dt}=\left.\frac{dF(t+\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0}=\left.\frac{d}{d\varepsilon}\left[F(t)F(\varepsilon)\right]\right|_{\varepsilon=0}=F(t)\left.\frac{dF(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} \tag{2.10}$$

Hence we have the identities

$$\dot F(t)=F(t)\dot F(0),\quad F(t)^{-1}\dot F(t)=\dot F(0) \tag{2.11}$$

The action of left multiplication by $F(t)^{-1}$ sends $\dot F(t)$ to $\dot F(0)=\text{const}\in T$.

$F(t)$ can be thought of as a "time evolution" operation, taking the velocity vector $\dot F(t_0)$ at some time t_0 to a future time $t_1>t_0$ by the action

$$\dot F(t_1)=F(t_1-t_0)\dot F(t_0)=F(t_1)F(-t_0)\dot F(t_0)=F(t_1)F(t_0)^{-1}\dot F(t_0)=F(t_1)\dot F(0). \tag{2.12}$$

Similarly, $F(t)^{-1}=F(-t)$ can be thought of as a time evolution backwards in time.

Just as we have $F(t)^{-1}\dot F(t)=\dot F(0)$, we in fact have a unique $F(t)$ such that

$$F(t)^{-1}\dot F(t)=A,\quad \forall A\in T. \tag{2.13}$$

If G is a matrix group[§] then we have the so-called **exponential map**

$$F(t)=\exp At \tag{2.14}$$

Suppose $F(t)\in U_0$ (the chart containing the identity) has local coordinates $f^i(t),\ i=1,...,n$. Since $F(t)$ is a one-parameter subgroup, its coordinate functions satisfy the same properties as the subgroup:

$$f^i(0)=0,\quad f^i(t_1+t_2)=\psi^i\left(f(t_1),f(t_2)\right),\quad f^i(-t)=\varphi^i(f(t)) \tag{2.15}$$

[§]Not all Lie groups have matrix representations, however by the Peter-Weyl Theorem every compact Lie group has a faithful finite-dimensional representation and is therefore isomorphic to a closed subgroup of $GL(n,\mathbb{C})$ for some n, i.e a matrix group.

where ψ and φ are the usual multiplication and inverse maps of the Lie algebra $\mathcal{G}$ of G.

If we consider left multiplication by $F(t)$, we have

$$x \to y = F(t)x, \quad y^i = \psi^i(f(t), x), \quad x, y \in G. \tag{2.16}$$

The velocity vector at $F(t)$ in terms of these coordinates is

$$\frac{dF}{dt} = \left(\dot{f}^1(t), ..., \dot{f}^n(t)\right), \quad \dot{f}^i(t) = \left.\frac{df^i(t+\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} = \left.\frac{d\psi^i(f(t), f(\varepsilon))}{d\varepsilon}\right|_{\varepsilon=0} = \left.\frac{\partial\psi^i(f(t), x)}{\partial x^j}\right|_{x=0} \dot{f}^j(0). \tag{2.17}$$

Since $F(t)$ is in the tangent space T, it induces a push-forward map

$$F_*(t) : \xi^i \to \eta^i = \frac{\partial\psi^i(f(t), x)}{\partial x^j}\xi^j, \quad \xi \in T_xG, \quad \eta \in T_yG. \tag{2.18}$$

Thus we see that $F_*(t)$ sends $\dot{F}(0)$ to $\dot{F}(t)$. Similarly, $F_*(t)^{-1}$ sends $\dot{F}(t)$ to $\dot{F}(0)$.

One-parameter subgroups can be used to define **canonical coordinates**[¶] in a neighbourhood of the identity of a Lie group G. We let $A_1, ..., A_n$ form a basis for the Lie algebra T. We know that for each

$$A = A_i x^i \in T \tag{2.19}$$

there is a one-parameter group $F(t) = \exp At$. We assign coordinates to the point $g = F(1) = \exp A$ which are the coefficients $x^1, ..., x^n$. This gives a system of coordinates in a sufficiently small neighbourhood of $g_0 = 1 \in G$. These are called the **canonical coordinates of the first kind**.

Of course, one might think that"first kind" implies the existence of a "second kind", and you'd be right. Another coordinate system can be constructed by introducing $F_i(t) = \exp A_i t$ and representing a point g sufficiently close to g_0 by

$$g = F_1(t_1)F_2(t_2)\cdots F_n(t_n) \tag{2.20}$$

for small t_i. Assigning the coordinates $x^1 = t_1, ..., x^n = t_n$ to the point g gives the **canonical coordinates of the second kind**.

Note that for a point g expressed in terms of canonical coordinates of the first and second kind

$$\begin{aligned} g_{(1)} &= \exp\left(A_1x^1 + A_2x^2 + ... + A_nx^n\right) \\ g_{(2)} &= \exp\left(A_1x^1\right)\exp\left(A_2x^2\right)\cdots\exp(A_nx^n) \end{aligned} \tag{2.21}$$

the two coordinate systems are equivalent only when $e^ae^b = e^{a+b}$, i.e. when $[A_i, A_j] = 0$.

2.3 Linear Representations

A **linear representation** of a group G is a group homomorphism

$$\rho : G \to GL(r, \mathbb{K}) \tag{2.22}$$

from G to the General Linear group of invertible $r \times r$ matrices such that

$$\rho(g_1g_2) = \rho(g_1)\rho(g_2) \quad \forall g_1, g_2 \in G. \tag{2.23}$$

Here, $\mathbb{K}$ is the field over which the matrices are defined (for our purposes, $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$).

Given a representation ρ of G, the map

$$\chi_\rho : G \to \mathbb{K}, \quad \chi_\rho(g) = \operatorname{tr}\rho(g), \quad g \in G \tag{2.24}$$

[¶] Also called Lie-Cartan coordinates

is called the **character** of the representation ρ.

A representation ρ of G is **irreducible** if the vector space $\mathbb{K}^r$ contains no proper subspaces invariant under the matrix group $\rho(G) = GL(r, \mathbb{K})$. Here, a subspace $W \subset \mathbb{K}^r$ is **invariant under the matrix group** $\rho(G)$ or simply G**-invariant** if

$$\rho(g)W \subset W \quad \forall g \in G. \tag{2.25}$$

We are then able to restrict ρ to W and get a **subrepresentation** of ρ.

Consider the representation

$$\rho : G \to GL(r_1, \mathbb{K}) \times GL(r_2, \mathbb{K}) \subset GL(r_1 + r_2, \mathbb{K}) \tag{2.26}$$

such that

$$g \to \rho(g) = \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix}, \quad \rho_k(g) \in GL(r_k, \mathbb{K}), \quad k = 1, 2. \tag{2.27}$$

The representation space is $\mathbb{K}^r = \mathbb{K}^{r_1} \oplus \mathbb{K}^{r_2}$ with elements

$$x = \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{K}^r, \quad \begin{pmatrix} u \\ 0 \end{pmatrix} \in \mathbb{K}^{r_1}, \quad \begin{pmatrix} 0 \\ v \end{pmatrix} \in \mathbb{K}^{r_2}. \tag{2.28}$$

We see that

$$\begin{aligned} \rho(g)\begin{pmatrix} u \\ 0 \end{pmatrix} &= \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix}\begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} \rho_1(g)u \\ 0 \end{pmatrix} \in \mathbb{K}^{r_1}, \\ \rho(g)\begin{pmatrix} 0 \\ v \end{pmatrix} &= \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ \rho_2(g)v \end{pmatrix} \in \mathbb{K}^{r_2}. \end{aligned} \tag{2.29}$$

Both $\mathbb{K}^{r_1}$ and $\mathbb{K}^{r_2}$ are G-invariant under ρ. Thus, we can write

$$\rho(g)x = \rho(g)\begin{pmatrix} u \\ v \end{pmatrix} = \rho(g)\begin{pmatrix} u \\ 0 \end{pmatrix} + \rho(g)\begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} \rho_1(g)u \\ \rho_2(g)v \end{pmatrix} = \rho_1(g)x + \rho_2(g)x. \tag{2.30}$$

ρ can be decomposed into a sum of representations $\rho = \rho_1 + \rho_2$. Such a representation is called a **reducible representation**.

Schur's Lemma:
Let

$$\rho_i : G \to GL(r_i, \mathbb{K}), \quad i = 1, 2 \tag{2.31}$$

be two irreducible representations (irreps) of G. If $A : \mathbb{K}^{r_1} \to \mathbb{K}^{r_2}$ is a linear transformation changing ρ_1 into ρ_2, i.e. A satisfies

$$A\rho_1(g) = \rho_2(g)A, \quad \forall g \in G, \tag{2.32}$$

then either A is the zero transformation or is a bijection, in which case $r_1 = r_2$.

If G is a Lie group and a representation $\rho : G \to GL(r, \mathbb{K})$ is a smooth map, then the push-forward map ρ_* is a linear map from the Lie algebra $\mathcal{G} = T_{(1)}$ to the space of $r \times r$ matrices:

$$\rho_* : \mathcal{G} \to M_{rr}(\mathbb{K}). \tag{2.33}$$

ρ_* is a representation of the Lie algebra $\mathcal{G}$, i.e. is a Lie algebra homomorphism:

1. it is linear
2. it preserves commutation: $\rho_*[\xi, \eta] = [\rho_*\xi, \rho_*\eta]$.

A representation is called **faithful**[‖] if it is one-to-one, i.e. $\rho(g) \neq \mathbb{1}$ unless $g = g_0$. If a Lie group has a faithful representation then it can be realised as a matrix Lie group, as mentioned in a previous footnote.

[‖]Mathematical monogamy joke.

2.4 The Adjoint Representation

An **inner automorphism** of a group G determined by some $h \in G$ is the transformation

$$\mathrm{AD} : G \to G, \quad g \to \mathrm{AD}(g) = hgh^{-1}, \quad g \in G. \tag{2.34}$$

Any inner automorphism does not move the identity $g_0 = hg_0h^{-1}$, and therefore the push-forward map of the tangent space $T = T_{g_0}G$ is a linear transformation of T denoted by $\mathrm{Ad}_h : T \to T$, satisfying

1. $\mathrm{Ad}_{g_0} = \mathbb{1}$, where $\mathbb{1}$ is the identity transformation of T
2. $\mathrm{Ad}_{h_1}\mathrm{Ad}_{h_2} = \mathrm{Ad}_{h_1h_2}$ $\forall h_1, h_2 \in G$, since $h_1h_2gh_2^{-1}h_1^{-1} = (h_1h_2)g(h_1h_2)^{-1}$
3. Choosing $h_1 = h$, $h_2 = h^{-1}$, we get $\mathrm{Ad}_{h^{-1}} = \mathrm{Ad}_h^{-1}$.

The map $h \to \mathrm{Ad}_h$ is a linear representation of G called the **adjoint**. For commutative Lie groups, the adjoint representation is the identity; $\mathrm{Ad}_h = 1$ $\forall h \in G$, since $hgh^{-1} = hh^{-1}g = g$.

In a neighbourhood of U_0, the push-forward map of the inner automorphism is

$$\mathrm{AD}(h)_* : \xi^i \to \eta^i = \frac{\partial \psi^i(\psi(h,x), \varphi(h))}{\partial x^j}\xi^j, \quad \xi \in T_gG, \quad \eta \in T_{hgh^{-1}}G. \tag{2.35}$$

If $x = 0$, then $g = g_0$ and $\xi, \eta \in T_{g_0}G = T$. The adjoint is then

$$\mathrm{Ad}_h : \xi \to \eta = \left.\frac{\partial \psi^i(\psi(h,x), \varphi(h))}{\partial x^j}\right|_{x=0} \xi^j = \left.\frac{\partial \psi^i(z, \varphi(h))}{\partial z^k}\right|_{z=h} \left.\frac{\partial \psi^k(h,x)}{\partial x^j}\right|_{x=0} \xi^j. \tag{2.36}$$

Let $F(t) = \exp At$ be a one-parameter subgroup of a Lie group G. Then $\mathrm{Ad}_{F(t)}$ is a one-parameter subgroup of $GL(n,\mathbb{R})$. The vector $\frac{d}{dt}\left.\mathrm{Ad}_{F(t)}\right|_{t=0}$ lies in the Lie algebra $gl(n,\mathbb{R})$ of $GL(n,\mathbb{R})$ and can be thought of as a linear operator denoted by ad_A:

$$\mathrm{ad}_A : \mathbb{R}^n \to \mathbb{R}^n, \quad B \to [A,B], \quad B \in T \cong \mathbb{R}^n. \tag{2.37}$$

A Lie algebra $\mathcal{G}$ is said to be

1. **simple** if it is noncommutative and has no proper ideals, i.e. subspaces $\mathcal{I} \subsetneq \mathcal{G}$ for which $[\mathcal{I}, \mathcal{G}] \subset \mathcal{I}$
2. **semisimple** if $\mathcal{G} = \mathcal{I}_1 \oplus \cdots \oplus \mathcal{I}_k$, where $\mathcal{I}_j$ are simple ideals with the property $[\mathcal{I}_i, \mathcal{I}_j] = 0$ for $i \neq j$.

The **Killing form** on a Lie algebra $\mathcal{G}$ is

$$\langle A, B \rangle = -\operatorname{tr}(\mathrm{ad}_A \mathrm{ad}_B). \tag{2.38}$$

If the Lie algebra $\mathcal{G}$ of a Lie group G is simple, then the linear representation $\mathrm{Ad} : G \to GL(n,\mathbb{R})$ is irreducible.

If the Killing form of $\mathcal{G}$ is positive definite, then the Lie algebra is semisimple.

3 Lie Derivatives and Covariant Differentiation

3.1 Vector and Tensor Fields

A **vector field** is a map that specifies a unique vector at each point x of a manifold M

$$\xi : M \to T(M), \quad x \to \xi_x \in T_x M. \tag{3.1}$$

A vector field intersects each tangent space T_x of the tangent bundle $T(M)$ at one and only one point.

In a coordinate basis (x^i_p), we can write

$$\xi = \xi^i_p(x) \frac{\partial}{\partial x^i_p}, \quad x \in U_p. \tag{3.2}$$

Since there is a *unique* vector at each point, we can drop the subscript p:

$$\xi = \xi^i(x) \frac{\partial}{\partial x^i}, \quad x \in M. \tag{3.3}$$

A vector field ξ can be understood as a differential operator that maps a scalar function to a scalar function on M

$$\xi(f) = \xi^i \frac{\partial f}{\partial x^i}. \tag{3.4}$$

This allows ξ to act as map from M to $\mathbb{R}$:

$$\xi(f) : M \to \mathbb{R}, \quad x \to \xi^i(x) \frac{\partial f(x)}{\partial x^i}. \tag{3.5}$$

Vector fields are linear and satisfy the Leibniz rule:

$$\xi(fg) = \xi(f)g + f\xi(g). \tag{3.6}$$

A **tensor field** of type (r, s) assigns a unique tensor of type (r, s) to each point $x \in M$.

$${}^{(r,s)}\xi : M \to T^{(r,s)}(M), \quad x \to^{(r,s)} \xi_x \in T^{(r,s)}_x M. \tag{3.7}$$

What happens if we compose vector fields? Consider the composition

$$\xi(\eta(f)) = \xi^i \frac{\partial}{\partial x^i}\left(\eta^j \frac{\partial f}{\partial x^j}\right) = \xi^i \frac{\partial \eta^j}{\partial x^i} \frac{\partial f}{\partial x^j} + \xi^i \eta^j \frac{\partial^2 f}{\partial x^i \partial x^j}. \tag{3.8}$$

Because of the second term, the composition is not a vector field. Instead, we can define the **Lie bracket**, or **commutator**

$$[\xi, \eta]f := \xi(\eta(f)) - \eta(\xi(f)) = \left(\xi^i \frac{\partial \eta^j}{\partial x^i} - \eta^i \frac{\partial \xi^j}{\partial x^i}\right) \frac{\partial f}{\partial x^j} \tag{3.9}$$

which is a vector field with components

$$[\xi, \eta]^j = \xi^i \frac{\partial \eta^j}{\partial x^i} - \eta^i \frac{\partial \xi^j}{\partial x^i}. \tag{3.10}$$

The Lie bracket has the properties

1. skew-symmetry: $[\xi, \eta] = -[\eta, \xi]$
2. linearity: $[\xi, \eta + \zeta] = [\xi, \eta] + [\xi, \zeta]$
3. $[\xi, f\eta] = f[\xi, \eta] + \xi(f)\eta$
4. Jacobi's identity: $[\xi, [\eta, \zeta]] + [\eta, [\zeta, \xi]] + [\zeta, [\xi, \eta]] = 0$.

Thus, a vector space of vector fields equipped with the commutator is an infinite dimensional Lie algebra.

3.2 Integral Curves

Let $\xi^i(x)$ be a vector field on M. Consider the autonomous (no explicit t dependence) system of differential equations

$$\dot{x}^i(t) := \frac{dx^i}{dt} = \xi^i(x^1(t), ..., x^n(t)), \quad i = 1, ..., n. \tag{3.11}$$

The solutions $x^i = x^i(t)$ to this system are called the **integral curves** of the vector field ξ^i and the vector field ξ^i itself is comprised of tangent vectors to the integral curves.

We can denote the integral curve of ξ^i by

$$F_t^i(x_0^1, ..., x_0^n) = x^i = x^i(t, x_0^1, ..., x_0^n), \quad x^i\big|_{t=t_0} = x_0^i. \tag{3.12}$$

Choosing $t_0 = 0$ with $x_0 := x_0^1, ..., x_0^n$, the formula $F_t^i(x_0) = x^i$ defines a self-map

$$F_t : (x_0) \to (x^1(t, x_0), ..., x^n(t, x_0)) \tag{3.13}$$

which depends on the parameter t.

In mechanics speak, F_t applied to the position $x_0 \in M$ is the time-evolution operator which gives the new position of a particle after time t as the particle moves along the integral curve through x_0.

Given a point $x_0 \in M$ with $(\xi^i) \neq 0$, the map F_t is locally a diffeomorphism, i.e. satisfies

$$F_{t+s} = F_t \circ F_s = F_s \circ F_t, \quad F_{-t} = F_t^{-1}. \tag{3.14}$$

The diffeomorphisms F_t define a **local group**. Needless to say, F_t is a one-parameter group.

This local abelian one-parameter group of diffeomorphisms F_t is called the **flow** generated by ξ^i. For small t, we can Taylor expand to find the solution

$$x^i(t, x_0) = x_0^i + t\xi^i(x_0) + \frac{1}{2}t^2 \frac{\partial \xi^i}{\partial x^j}\xi^j(x_0) + \mathcal{O}(t^3). \tag{3.15}$$

For some one-parameter local group of diffeomorphisms $F_t = (F_t^1, ..., F_t^n)$, we can define its **velocity field** to be the vector field

$$\xi^i = \frac{d}{dt}F_t^i\bigg|_{t=0}, \quad i = 1, ..., n. \tag{3.16}$$

The commutator $[\xi, \eta]$ can be interpreted geometrically as the measure of the discrepancy between the points arrived at by following the integral curves ξ and η (with flows F_t and G_s, respectively) in different orders. For small t and s, we have (after arduous calculation)

$$[G_s, F_t](x) = G_s(F_t(x)) - F_t(G_s(x)) = ts[\xi, \eta] + \mathcal{O}(t^3, s^3). \tag{3.17}$$

If the Lie bracket of two vector fields vanishes, the vector fields **commute**. Note that vectors comprising a coordinate induced basis commute because partial derivatives commute. The converse is also true; if all the elements of a basis for vector fields commute then the basis is coordinate induced.

A one-parameter group of diffeomorphisms $F_t(x)$ with vector field $\xi(x)$ acts on smooth functions $f = f(x)$ as

$$(F_t f)(x) = f(F_t(x)). \tag{3.18}$$

e.g. the one-parameter group of translations $F_t(x) = x + t$ with $\xi = 1$ acts as

$$F_t f(x) = f(x + t). \tag{3.19}$$

For a general analytic function $f(x)$, we have

$$\begin{aligned} F_t f(x) = f(x+t) &= f(x) + tf'(x) + \frac{1}{2}t^2 f''(x) + ... \\ &= \left(\sum_{k=0}^{\infty} \frac{t^k}{k!}\frac{d^k}{dt^k}\right) f(x) := \exp\left(t\frac{d}{dt}\right) f(x) = e^{t\partial_\xi} f(x). \end{aligned} \tag{3.20}$$

Here, $\exp(t\partial_\xi)$ is the **exponential function** of ξ. $\partial_\xi = \xi^i \dfrac{\partial}{\partial x^i}$ is the directional derivative in the direction of ξ.

WE can use this to define an action of the flow F_t generated by the vector field ξ on a tensor $T = (T^{(i)}_{(j)})$ of type (p,q) where $(i) = i_1, \cdots i_p$, $(j) = j_1 \cdots j_q$. We consider a region on which F_t are one-to-one

$$\dot{x}^i = \xi^i(x) \implies y^i(t) = F^i_t(x), \quad y^i(0) = x^i. \tag{3.21}$$

Since x and y are different points on the manifold, we cannot compare them. Instead, we perform a passive transformation of the basis vectors by F_t^{-1}. This moves the point $y = F_t(x)$ to the point x. We then define the action of $F_t(x)$ on T by

$$(F_tT)^{(i)}_{(j)}(x) = T^{(k)}_{(l)}(y)\frac{\partial y^{l_1}}{\partial x^{j_1}}\cdots\frac{\partial y^{l_q}}{\partial x^{j_q}}\frac{\partial x^{i_1}}{\partial y^{k_1}}\cdots\frac{\partial x^{i_p}}{\partial y^{k_p}}. \tag{3.22}$$

We see that $F_tT(x)$ and $T(y)$ are simply the same tensor measured in different coordinate systems.

We can now define the **Lie derivative** of a tensor $T = (T^{(i)}_{(j)})$ along a vector field ξ as the tensor $L_\xi T$ given by

$$L_\xi T^{(i)}_{(j)} = \frac{d}{dt}(F_tT)^{(i)}_{(j)}\bigg|_{t=0}. \tag{3.23}$$

If we regard F_t as a time-dependent deformation of the manifold, then the Lie derivative measures the rate of change of the tensor T resulting from this deformation.

The explicit formula for the Lie derivative is

$$\begin{aligned} L_\xi T^{(i)}_{(j)} = \xi^\alpha \frac{\partial T^{(i)}_{(j)}}{\partial x^\alpha} &+ T^{i_1\cdots i_p}_{\alpha j_2\cdots j_q}\frac{\partial \xi^\alpha}{\partial x^{j_1}} + ... + T^{i_1\cdots i_p}_{j_1\cdots j_{q-1}\alpha}\frac{\partial \xi^\alpha}{\partial x^{j_q}} \\ &- T^{\alpha i_2\cdots i_p}_{j_1\cdots j_q}\frac{\partial \xi^{i_1}}{\partial x^\alpha} - ... - T^{i_1\cdots i_{p-1}\alpha}_{j_1\cdots j_q}\frac{\partial \xi^{i_p}}{\partial x^\alpha}. \end{aligned} \tag{3.24}$$

If $T = f$ is a scalar, then the Lie derivative is the directional derivative:

$$L_\xi f = \xi^\alpha \frac{\partial f}{\partial x^\alpha} = \partial_\xi f. \tag{3.25}$$

If $L_\xi f = 0$, then f is constant along the integral curves of ξ and f is called an **integral of the field**. e.g. for $\xi = (-y, x)$ the functions $f(x,y) = x^2 + y^2 - c$ are integral.

If $T = \eta$ is a vector field, the Lie derivative is

$$L_\xi \eta^i = \xi^\alpha \frac{\partial \eta^i}{\partial x^\alpha} - \eta^\alpha \frac{\partial \xi^i}{\partial x^\alpha} = [\xi, \eta]^i \tag{3.26}$$

and has the properties

1. $L_\xi \eta = -L_\eta \xi$

2. $L\xi(f\eta) = fL\xi\eta + \eta\partial_\xi f$

3. $L_{[\xi,\eta]}f = \partial_{[\xi,\eta]} = [\partial_\xi, \partial_\eta]f = [L_\xi, L_\eta]f$.

If $T = (T_i)$ is a covector, then the Lie derivative is

$$L_\xi T_i = \xi^\alpha \frac{\partial T_i}{\partial x^\alpha} + T_\alpha \frac{\partial \xi^\alpha}{\partial x^i}. \tag{3.27}$$

In particular, if $T_i = \dfrac{\partial f}{\partial x^i} := df_i$, then

$$L_\xi df_i = \xi^\alpha \frac{\partial^2 f}{\partial x^\alpha \partial x^i} + \frac{\partial f}{\partial x\alpha}\frac{\partial \xi^\alpha}{\partial x^i} = \frac{\partial}{\partial x^i}\left(\xi^\alpha \frac{\partial f}{\partial x^\alpha}\right) = \frac{\partial}{\partial x^i}(L_\xi f). \tag{3.28}$$

So, L_ξ and d commute; $L_\xi df = d(L_\xi f)$.

If $T = (g_{ij})$ is a tensor of type $(0, 2)$, then

$$L_\xi g_{ij} = \xi^\alpha \frac{\partial g_{ij}}{\partial x^\alpha} + g_{\alpha j}\frac{\partial \xi^\alpha}{\partial x^i} + g_{i\alpha}\frac{\partial \xi^\alpha}{\partial x^j} := u_{ij}, \tag{3.29}$$

where u_{ij} is called the **strain tensor**. If g_{ij} is a metric tensor of M then u_{ij} describes hoe g_{ij} changes under small deformations F_t defined by ξ.

If the space is Euclidean, then $g_{ij} = \delta_{ij}$ and

$$u_{ij} = \frac{\partial \xi^i}{\partial x^i} + \frac{\partial \xi^j}{\partial x^j} \tag{3.30}$$

If $L_\xi g_{ij} = 0$, then ξ is called a **Killing vector**. The metric g_{ij} does not change under deformations F_t defined by the Killing vector field ξ.

In a coordinate system (y^i) where a Killing vector field $\xi = (1, 0, ..., 0)$, then the metric g_{ij} is independent of y^1.

If ξ and η are Killing vectors, then their commutator is also a Killing vector:

$$L_{[\xi,\eta]}g_{ij} = [L_\xi, L_\eta]g_{ij} = 0. \tag{3.31}$$

So, the Killing vector fields of a pseudo-Riemannian manifold form a Lie algebra with respect to the Lie bracket.

If $T = \sqrt{|g|}\varepsilon_{i_1\cdots i_n}$ is the volume element, then

$$L_\xi \sqrt{|g|}\varepsilon_{i_1\cdots i_n} = \varepsilon_{i_1\cdots i_n}\frac{\partial}{\partial x^\alpha}\left[\sqrt{|g|}\xi^\alpha\right]. \tag{3.32}$$

The volume of an oriented, closed manifold does not change under small deformations F_t defined by ξ.

The quantity

$$\nabla_\alpha \xi^\alpha := \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^\alpha}\left[\sqrt{|g|}\xi^\alpha\right] \tag{3.33}$$

is a scalar called the **divergence** of the vector field ξ. Note that

$$\nabla_\alpha \xi^\alpha = \frac{1}{2}g^{ij}u_{ij} = \frac{1}{2}g^{ij}L_\xi g_{ij}. \tag{3.34}$$

3.3 Covariant Differentiation

The differential d transforms a skew-symmetric tensor T into another skew-symmetric tensor dT. e.g.

$$(dT)_{ij} = \frac{\partial T_j}{\partial x^i} - \frac{\partial T_i}{\partial x^j} = \partial_i T_j - \partial_j T_i. \tag{3.35}$$

It is important that this quantity dT is a tensor, because $\partial_j T_i$ alone is not a tensor. We can see this by changing coordinates.

$$\begin{aligned}\partial_q \widetilde{T}_j = \frac{\partial \widetilde{T}_j}{\partial z^q} &= \frac{\partial}{\partial z^q}\left(T_i \frac{\partial x^i}{\partial z^j}\right) \\ &= \frac{\partial T_i}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i \frac{\partial^2 x^i}{\partial z^q \partial z^j} \\ &= \frac{\partial T_i}{\partial x^p}\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i \frac{\partial^2 x^i}{\partial z^q \partial z^j}.\end{aligned} \tag{3.36}$$

The first term is the usual tensor transformation rule, however the second term being non-zero means that this quantity is not a tensor. What we would like is a tensorial version of the partial derivative. We can do this by introducing an operator ∇_k - the **covariant derivative** - such that in Euclidean coordinates

$$\nabla_k T^{(i)}_{(j)} = \frac{\partial T^{(i)}_{(j)}}{\partial x^k}. \tag{3.37}$$

For this to be a tensor, we impose that it must transform like a tensor. i.e.,

$$\nabla_r \widetilde{T}^{(k)}_{(l)} = \nabla_s T^{(i)}_{(j)} \frac{\partial x^s}{\partial z^r}\frac{\partial x^{(j)}}{\partial z^{(l)}}\frac{\partial z^{(k)}}{\partial x^{(i)}}. \tag{3.38}$$

We can also write this operator using semicolon notation: $\nabla_k T^i_j := T^i_{j;k}$.

Consider the vector T^i. Using $\widetilde{T}^k = T^i \frac{\partial z^k}{\partial x^i}$ and $T^i = \widetilde{T}^k \frac{\partial x^i}{\partial z^k}$, we have

$$\begin{aligned}\nabla_r \widetilde{T}^k &= \partial_s T^i \frac{\partial x^s}{\partial z^r}\frac{\partial z^k}{\partial x^i} \\ &= \frac{\partial T^i}{\partial x^s}\frac{\partial x^s}{\partial z^r}\frac{\partial z^k}{\partial x^i} \\ &= \frac{\partial T^i}{\partial z^r}\frac{\partial z^k}{\partial x^i} \\ &= \frac{\partial}{\partial z^r}\left(T^i \frac{\partial z^k}{\partial x^i}\right) - T^i \frac{\partial}{\partial z^r}\frac{\partial z^k}{\partial x^i} \\ &= \frac{\partial \widetilde{T}^k}{\partial z^r} - \widetilde{T}^s \frac{\partial x^i}{\partial z^s}\frac{\partial x^m}{\partial z^r}\frac{\partial^2 z^k}{\partial x^m \partial x^i}\end{aligned} \tag{3.39}$$

Introducing the **connection coefficients**

$$\Gamma^k_{sr} = -\frac{\partial x^i}{\partial z^s}\frac{\partial x^m}{\partial z^r}\frac{\partial^2 z^k}{\partial x^m \partial x^i}, \tag{3.40}$$

we have the equation for the covariant derivative of a vector;

$$\nabla_r \widetilde{T}^k = \frac{\partial \widetilde{T}^k}{\partial z^r} + \Gamma^k_{sr}\widetilde{T}^s. \tag{3.41}$$

The formula for a covector is pretty much the same, just with a minus sign on Γ:

$$\nabla_r \widetilde{T}_k = \frac{\partial \widetilde{T}_k}{\partial z^r} - \Gamma^s_{kr}\widetilde{T}_s. \tag{3.42}$$

For an arbitrary rank tensor, the process is is simply to add a set of connection coefficients for each upper index and subtract one for each lower index. i.e. for a (p, q) tensor

$$\nabla_r T^{(k)}_{(l)} = \frac{\partial T^{(k)}_{(l)}}{\partial z^r} + \sum_{\alpha=1}^{p} \Gamma^{k_\alpha}_{sr} T^{k_1 \cdots (k_\alpha = s) \cdots k_p}_{l_1 \cdots l_q} - \sum_{\alpha=1}^{q} \Gamma^{s}_{l_\alpha r} T^{k_1 \cdots k_p}_{l_1 \cdots (l_\alpha = s) \cdots l_q}. \tag{3.43}$$

A connection is called **affine** or **Euclidean** if there exists a coordinate system where $\Gamma^k_{ij} = 0$. Such coordinates are also called affine.

Note that these connection coefficients are not tensors. Their purpose is to facilitate covariant differentiation being a tensor, as a correcting term. However, the quantity

$$T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj} = \Gamma^i_{[jk]} \tag{3.44}$$

is a tensor, called the **torsion tensor**. A connection is **torsion-free** or **symmetric** if

$$T^i_{jk} = 0 \iff \Gamma^i_{jk} = \Gamma^i_{kj}. \tag{3.45}$$

Covariant differentiation has some important properties:

1. It is linear and commutes with contraction of indices.
2. The covariant derivative of a product is computed using the Leibnitz product rule:

$$\nabla_k (T^{(i)}_{(p)} S^{(j)}_{(q)}) = (\nabla_k T^{(i)}_{(p)}) S^{(j)}_{(q)} + T^{(i)}_{(p)} (\nabla_k S^{(j)}_{(q)}). \tag{3.46}$$

3. The covariant derivative of the basis vectors and basis covectors can be thought of as defining the connection coefficents:

$$\nabla_k e_i = \Gamma^j_{ik} e_j, \quad \nabla_k e^i = -\Gamma^i_{jk} e^j. \tag{3.47}$$

4 Parallel Transport and Geodesics

4.1 Parallel Transport of Tensor Fields

The **directional derivative** of a (p,q) tensor $T = (T^{(i)}_{(j)})$ at the point $P \in M$ along the vector $\xi \in T_P M$ is the (p,q) tensor

$$\nabla_\xi T^{(i)}_{(j)} = \xi^k \nabla_k T^{(i)}_{(j)}. \tag{4.1}$$

For a scalar f, this definition coincides with the usual definition of the directional derivative of a function,

$$\nabla_\xi f = \xi^k \partial_k f = \partial_\xi f. \tag{4.2}$$

Normally for a directional derivative of a scalar, we can let $\xi(t)$ be a velocity vector of a curve $\mathcal{C}$, i.e.

$$x^i = x^i(t), \quad \xi^i(t) = \frac{dx^i}{dt}, \quad i = 1, ..., n. \tag{4.3}$$

If we have $\partial_\xi f = 0$ for all points on $\mathcal{C}$, then we say the function f is constant on $\mathcal{C}$.

A reasonable question to ask might be whether we can say that two different vectors or tensors at different points are parallel to one another. In flat space this is trivial; we move one vector over to the other and see of they point in the same direction.
If space is curved, however, this is not so simple. In general, there is no unique way to move a vector or tensor from one point to another, and no one way is better than any other.
This is a big problem, and has no solution. We must simply give up the comfortable idea of comparing vectors in curved space unless they are both in the same tangent space.
This has some real-world implications. For example, in curved space there is no definite notion of relative velocity - the concept is meaningless since we can't compare the two velocity vectors. In cosmology, the gravitational redshift experienced by galaxies and other objects *seems* like the regular Doppler shift; the equations are even of the same form. However, ascribing this frequency shift to relative motion makes no sense. In some cases it even leads to galaxies receding from one another at faster than the speed of light, which is of course nonsense. Locally, neither object is actually moving that fast. Anyway, back to the maths.

We can use our covariant derivative to define constancy along a curve. We say that a tensor field T is **covariantly constant** or **parallel** along a curve $\mathcal{C}$ if

$$\nabla_\xi T = \xi^k \nabla_k T = 0, \quad \xi^k = \frac{dx^k}{dt}, \quad a \leq t \leq b. \tag{4.4}$$

In the case of a vector field, we have

$$\begin{aligned} \nabla_\xi T^i = \xi^k \nabla_k T^i &= \frac{dx^k}{dt}\left(\frac{\partial T^i}{\partial x^k} + \Gamma^i_{jk} T^j\right) \\ &= \frac{dT^i(x(t))}{dt} + \frac{dx^k}{dt}\Gamma^i_{jk} T^j = 0. \end{aligned} \tag{4.5}$$

Equation (4.5) is called the **equation of parallel transport**. This definition of parallelism is coordinate independent because covariant differentiation is a tensor operation. Note as well that it is dependent on the given connection.

In general, since parallelism depends on the curve $\mathcal{C}$, there is no guarantee that a tensor will be constant along another, different curve. However, if the connection is affine, $\Gamma^i_{jk} = 0$, then we can have tensors which are covariantly constant and non-zero everywhere on a chart U.

4.2 Geodesics

Geodesics are a curved-space generalisation of the notion of "straight lines". There's many ways of defining what we mean by a straight line, the most common perhaps being that it's the shortest path between two points. Of course, this doesn't work well for us because we're working with individual points and it's tricky to compare them. A more practical definition is that a straight line parallel transports its own tangent vector. That is, the vector tangent to a straight line is constant; it doesn't "change direction". Let's see what this looks like.

For a curve $x^i = x^i(t)$ parametrised by t, its tangent vector is $\dot{x}^i(t) = \dfrac{dx^i}{dt}$. The curve $x^i(t)$ is called a **geodesic** if

$$\nabla_{\dot{x}}(\dot{x}) = 0. \tag{4.6}$$

Explicitly, we get

$$\nabla_{\dot{x}}(\dot{x}) = \frac{dx^k}{dt}\nabla_k \frac{dx^i}{dt} = \frac{d}{dt}\frac{dx^i}{dt} + \frac{dx^k}{dt}\Gamma^i_{jk}\frac{dx^j}{dt} = 0, \tag{4.7}$$

so geodesic curves satisfy the system of n second order differential equations

$$\frac{d^2x^i}{dt^2} + \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0, \quad i = 1, ..., n. \tag{4.8}$$

This is called the **geodesic equation** and has a unique solution for given initial conditions $x^i(0) = x^i_0$ and $\dot{x}^i(0) = v^i_0$.

We can see immediately that if we're in an affine - or Euclidean - space, where $\Gamma^i_{jk} = 0$, then the geodesic equation is just

$$\frac{d^2x^i}{dt^2} = 0, \tag{4.9}$$

which has solutions $x^i(t) = v^i_0 t + x^i_0$, the equation of a straight line that we all know and love.

It's worth noting that there is not always a single geodesic. For example, two points on the equator of a sphere $\mathcal{S}^2$ has two geodesics - the shorter path along the equator and the longer path going around the equator the other way.

4.3 Metric-Compatibility and the Connection

An important property of parallel transport arises when our manifold has a metric structure on it. Namely, if the covariant derivative is **metric-compatible**, i.e. $\nabla_k g_{ij} = 0$, and two vector fields T^i and S^j are both parallel along the curve, then parallel transport preserves their inner product:

$$\frac{d}{dt}\langle T, S\rangle = \frac{d}{dt}\left[g_{ij}T^iS^j\right] = \frac{dx^k}{dt}\nabla_k(g_{ij}T^iS^j) = g_{ij}\frac{dx^k}{dt}\nabla_k(T^iS^j) = 0. \tag{4.10}$$

This is an important property of parallel transport. It means that it preserves vector norms (it doesn't randomly alter the lengths of vectors) and it's also an orthogonal transformation (the angle between two vectors in the tangent space is fixed under parallel transport).

The geodesic equation we have above is very nice and exists perfectly well on its own, but we would like to be able to connect it to the other structures we have defined on our manifold. Specifically, when the manifold has a metric defined on it.

We can define a unique connection on a manifold with metric g_{ij} by imposing two conditions:

1. The connection is torsion-free: $\Gamma^k_{ij} = \Gamma^k_{ji}$
2. The covariant derivative with respect to Γ^k_{ij} is metric-compatible: $\nabla_k g_{ij} = 0$.

These conditions give a unique form of the connection, called the **Christoffel connection****, whose components

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} \left(\partial_i g_{lj} + \partial_k g_{il} - \partial_l g_{ij} \right) \tag{4.11}$$

are called the **Christoffel components** or **Christoffel symbols**. It's important to note that this connection is unique and always exists where we can define a metric.

4.4 The Curvature Tensor

As we have seen, parallel transport of a vector ξ is determined by the equation of parallel transport. If $\Gamma^i_{jk} = 0$, this equation is trivial to solve, so it would be nice to have a way to find out if these connection coefficients are zero. i.e. which manifolds are flat. What we need, then, is a way to measure the curvature of a manifold.

If $\Gamma^i_{jk} = 0$, then we know the covariant derivative coincides with the partial derivative;

$$\nabla_k T^{(i)}_{(j)} = \partial_k T^{(i)}_{(j)}. \tag{4.12}$$

Since partial derivatives commute, if the connection coefficients are zero we have

$$[\nabla_k, \nabla_l] = (nabla_k \nabla_l - \nabla_l \nabla_k) T^{(i)}_{(j)} = 0. \tag{4.13}$$

In this case however, the left hand side of this equation is a tensor, so it must be true in any coordinate system. This means that if we have an arbitrary connection, if $[\nabla_k, \nabla_l]$ with respect to that connection is zero, the connection is Euclidean. We can expand this equation explicitly to find

$$[\nabla_k, \nabla_l]\xi^i = (\partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} + \Gamma^m_{jl}\Gamma^i_{mk} - \Gamma^m_{jk}\Gamma^i_{ml})\xi^j + (\Gamma^j_{kl} - \Gamma^j_{lk})\nabla_j \xi^i \tag{4.14}$$

The first quantity on the right hand side is given the notation

$$R^i_{jkl} = \partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} + \Gamma^m_{jl}\Gamma^i_{mk} - \Gamma^m_{jk}\Gamma^i_{ml} \tag{4.15}$$

and is called the **Riemann curvature tensor**. The other term is the familiar torsion tensor, T^j_{kl}. This gives us the **Ricci identity**

$$[\nabla_k, \nabla_l]\xi^i = R^i_{jkl}\xi^j + T^j_{kl}\nabla_j \xi^i. \tag{4.16}$$

For a symmetric, torsion-free connection, if the Riemann curvature tensor is identically zero then the connection is Euclidean.

The curvature tensor has some important symmetries which can be derived from its explicit form.

1. It is skew-symmetric in its last two indices: $R^i_{jkl} + R^i_{jlk} = 0$.
2. If the connection is symmetric, then $R^i_{jkl} + R^i_{klj} + R^i_{ljk} = 0$.
3. If the connection is metric-compatible and we define the purely covariant curvature tensor

   $$R_{ijkl} = g_{im} R^m_{jkl}, \tag{4.17}$$

   then R_{ijkl} is skew-symmetric in the first two indices: $R_{ijkl} + R_{jikl} = 0$.
4. If the connection is symmetric and metric-compatible, then R_{ijkl} is symmetric under exchange of the first two indices with the last two: $R_{ijkl} = R_{klij}$.

**This connection goes by many names, including but not limited to the Christoffel connection, the Levi-Civita connection, and the Riemannian connection.

4.5 Normal Coordinates

A useful form of the metric is the **canonical form**

$$g_{ij} = \mathrm{diag}(-1, -1, ..., -1, +1, +1, ..., +1, 0, ..., 0). \tag{4.18}$$

If the manifold has dimension n and the canonical form has t (-1)'s and s $(+1)$'s, then $s - t$ is the metric **signature** and $s + t$ is the metric **rank**. If the metric is continuous, it has the same rank everywhere. If the rank is equal to the dimension of the manifold, it is non-degenerate.

It is always possible to put the metric into its canonical form at a single point $P \in M$, but not necessarily in a neighbourhood of P and certainly not over the whole manifold (unless the manifold is flat). At any point P, there exists a coordinate system such that g_{ij} takes its canonical form and the derivatives $\partial_k g_{ij}$ are all zero. These coordinates are called **Riemann normal coordinates (RNCs)**, and the basis vectors of this coordinate system form a **local Lorentz frame**. This is a formal statement of the idea behind manifolds, that they "look locally like flat space" up to first order. Note that in RNCs, the second derivatives $\partial_l \partial_k g_{ij}$ do not usually vanish.

Of course, since the connection is made up of the metric and its first derivatives, $\Gamma^k_{ij} = 0$ at the point P.

4.6 Contractions of the Curvature Tensor

Our curvature tensor R^i_{jkl} has built into it the information about the curvature of its manifold with respect to some connection. We discussed some of the symmetries of this tensor, each of which reduces its number of independent components. Cutting to the chase, the number of independent components in n dimensions is

$$\frac{1}{12} n^2 (n^2 - 1). \tag{4.19}$$

In 1, 2, 3, and 4 dimensions, there are 0, 1, 6, and 20 independent components. There are some important takeaways to this. Firstly, the 20 independent functions in four dimensions correspond exactly to the 20 degrees of freedom in the second derivatives of the metric which we could not set to zero in RNCs.

Secondly, notice that in 1 dimension there are no independent components of the curvature tensor. One dimensional manifolds such as $\mathcal{S}^1$ - the circle - are always flat. We think of a circle as having curvature because it is embedded in 2d space. This is called **extrinsic curvature**.

Now, sometimes it is useful to consider other tensors built from the curvature tensor. For the Christoffel connection, there is one independent contraction of R^i_{jkl} called the **Ricci tensor**, formed by contracting the upper index with the second lower index

$$R_{ij} = R^k_{ikj}. \tag{4.20}$$

If we have the metric, we can contract this tensor again to get the **Ricci scalar**, or **scalar curvature**

$$R = R^k{}_k = g^{jk} R_{jk}. \tag{4.21}$$

In $n = 2$ dimensions, where the curvature tensor has only one independent component, the Ricci scalar contains all the information about the curvature of the space.

We can construct an important tensor from the Ricci tensor and scalar curvature. Using the **Bianchi identity** (without proof)

$$\nabla_{[m} R_{ij]kl} = 0, \tag{4.22}$$

and contracting twice, we find that

$$\nabla^k R_{jk} = \frac{1}{2} \nabla_j R. \tag{4.23}$$

This means that if we construct the **Einstein tensor**

$$G_{ij} = R_{ij} - \frac{1}{2}Rg_{ij}, \tag{4.24}$$

we have

$$\nabla^i G_{ij} = 0. \tag{4.25}$$

The Einstein tensor appears in the Einstein field equations

$$G_{ij} + \Lambda g_{ij} = \frac{8\pi G}{c^4}T_{ij} \tag{4.26}$$

where Λ is the cosmological constant, G is Newton's gravitational constant, and T_{ij} is the stress-energy or energy-momentum tensor. This equation was Einstein's 1915 masterpiece; the left hand side encapsulates the curvature of spacetime through the metric, and the right hand side describes the matter and energy distribution of space. i.e. matter and energy curve spacetime.

The Ricci tensor and scalar curvature tell us about the traces of the Riemann curvature tensor. Of course, there is other information in there too. We can construct the **Weyl tensor**, which is the curvature tensor with all of its contractions subtracted off.

$$C_{ijkl} = R_{ijkl} - \frac{2}{n-2}\left(g_{i[k}R_{l]j} - g_{j[k}R_{l]i}\right) + \frac{2}{(n-2)(n-1)}g_{i[k}g_{l]j}R. \tag{4.27}$$

The Weyl tensor is sometimes called the conformal tensor because it is invariant under conformal transformations

$$g_{ij} \to \lambda(x)g_{ij}, \quad \lambda(x) \in C^\infty(M). \tag{4.28}$$

5 Appendix

This appendix contains some extra information primarily based on discussions from Sean Carroll's *Lecture Notes on General Relativity*, which I have already mentioned.

5.1 The Parallel Propagator

We have seen already the equation of parallel transport

$$\frac{dT^i(x(t))}{dt} + \frac{dx^k}{dt}\Gamma^i_{jk}T^j = 0. \tag{5.1}$$

We can, in fact, find an explicit, general solution to this equation. For some path $x(t)$, solving the equation of parallel transport amounts to finding an operator $P(t, t_0)$ which takes a vector $T^i(t_0)$ from an initial point at t_0 to a final value t further along the path, given by

$$T^i(t) = P^i_j(t, t_0)T^j(t_0). \tag{5.2}$$

The matrix operator $P^i_j(t, t_0)$ is called the **parallel propagator**. If we define $A^i_k(t) = -\Gamma^i_{jk}\frac{dx^j}{dt}$, then the equation for parallel transport becomes

$$\frac{dT^i}{dt} = A^i_k T^k. \tag{5.3}$$

Substituting in for (5.2), using the product rule, we get

$$\frac{d}{dt}P^i_k(t, t_0) = A^i_j(t)P^j_k(t, t_0). \tag{5.4}$$

We can solve this equation by integrating both sides;

$$P^i_k(t, t_0) = \delta^i_k + \int_{t_0}^{t} A^i_j(t')P^j_k(t', t_0)dt'. \tag{5.5}$$

Now, taking this equation and plugging it into itself, we get

$$P^i_k(t, t_0) = \delta^i_k + \int_{t_0}^{t} A^i_k(t')dt' + \int_{t_0}^{t}\int_{t_0}^{t'} A^i_j(t')A^j_k(t'')dt''dt' + ... \tag{5.6}$$

The n^{th} term in this sequence is an integral over an n-dimensional right triangle, called an n-simplex. Instead of integrating over a simplex, we can integrate over a hypercube, which contains $n!$ simplexes, so we multiply the n^{th} integral by $\frac{1}{n!}$. We also need the integrand $A(t_n)A(t_{n-1})\cdots A(t_1)$ to be ordered such that $t_n \geq t_{n-1} \geq \cdots \geq t_1$. We do this by introducing the **path ordering operator** $\mathcal{P}$. Our general term is then

$$\frac{1}{n!}\int_{t_0}^{t}\cdots\int_{t_0}^{t} \mathcal{P}\left[A(t_n)A(t_{n-1})\cdots A(t_1)\right] d^n t. \tag{5.7}$$

We can now write our parallel propagator as

$$P(t, t_0) = 1 + \sum_{n=1}^{\infty}\frac{1}{n!}\int_{t_0}^{t} \mathcal{P}\left[A(t_n)A(t_{n-1})\cdots A(t_1)\right] d^n t. \tag{5.8}$$

This is the series expansion of an exponential, so we can say that the parallel propagator is given by the path-ordered exponential

$$\begin{aligned} P(t, t_0) &= \mathcal{P}\exp\left(\int_{t_0}^{t} A(t)dt\right) \\ P^i_j(t, t_0) &= \mathcal{P}\exp\left(-\int_{t_0}^{t} \Gamma^i_{kj}\frac{dx^k}{dt}dt\right). \end{aligned} \tag{5.9}$$

This same kind of expression appears in quantum perturbation theory as Dyson's formula, because the Schrödinger equation has the same form as equation (5.4).

For a metric-compatible connection, if the path followed is a loop, starting and ending at the same point, then the parallel propagator is just a Lorentz transformation of the tangent space. The transformation is called the **holonomy** of the loop. Knowing the holonomy of all possible loops is equivalent to knowing the metric, which led to research into describing general relativity in terms of these loops; Loop Quantum Gravity.

5.2 Geodesics and the Christoffel Connection

In a Lorentzian spacetime, which has a pseudo-Riemann metric g_{ij}, we have three different types of path; space-like, time-like, and light-like (or null), each with different path lengths described by the metric as

$$ds^2 = g_{ij} dx^i dx^j \tag{5.10}$$

For space-like paths, $ds^2 > 0$, for time-like paths $ds^2 < 0$, and for null paths $ds^2 = 0$. A useful parametrisation for time-like paths is the proper time τ, which can be thought of as the total time measured by a clock moving along the curve.

$$\tau = \int \sqrt{-g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}}dt \tag{5.11}$$

One of our definitions for straight lines was that they are the shortest distance between two points. This corresponds to finding the extrema of the proper time (spoiler alert: the shortest paths will be the maxima of the proper time).

We perform calculus of variations on τ in order to find the extrema. We thus consider an infinitesimal change

$$\begin{aligned} x^i &\to x^i + \delta x^i \\ g_{ij} &\to g_{ij} + \delta x^k \partial_k g_{ij} \end{aligned} \tag{5.12}$$

The change in the metric comes from considering its Taylor expansion. Putting this into our τ equation, we get

$$\begin{aligned} \tau + \delta\tau &= \int \left(-g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} - \partial_k g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\delta x^k - 2g_{ij}\frac{dx^i}{dt}\frac{d(\delta x^j)}{dt}\right)^{1/2} dt \\ &= \int \left(-g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right)^{1/2}\left[1 + \left(-g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right)^{-1} \times \right. \\ &\qquad \left. \times \left(-\partial_k g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\delta x^k - 2g_{ij}\frac{dx^i}{dt}\frac{d(\delta x^j)}{dt}\right)\right]^{1/2} dt. \end{aligned} \tag{5.13}$$

Since δx^k is small, we can expand the term in square brackets. The 0^{th} order term cancels τ (try it) and we get

$$\delta\tau = \int \left(-g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right)^{-1/2}\left(-\frac{1}{2}\partial_k g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\delta x^k - g_{ij}\frac{dx^i}{dt}\frac{d(\delta x^j)}{dt}\right) dt. \tag{5.14}$$

We can remove the leading factor in this expression by changing the parameter of our curve $x(t)$ from t to the proper time itself using

$$dt = \left(-g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right)^{-1/2} d\tau. \tag{5.15}$$

We substitute in for dt (everywhere, including in the derivative terms) and get

$$\delta\tau = \int \left(-\frac{1}{2}\partial_k g_{ij} \frac{dx^i}{d\tau}\frac{dx^j}{d\tau}\delta x^k - g_{ij}\frac{dx^i}{d\tau}\frac{d(\delta x^j)}{d\tau} \right) d\tau = \int \left(-\frac{1}{2}\partial_k g_{ij} \frac{dx^i}{d\tau}\frac{dx^j}{d\tau} + \frac{d}{d\tau}\left[g_{ik}\frac{dx^i}{d\tau} \right] \right) \delta x^k d\tau. \tag{5.16}$$

In the second line we integrated by parts and demanding that δx^k vanish at the boundary points. Now, we can find conditions for the extrema of τ. We want $\delta\tau$ to vanish for arbitrary variations δx^k, so we need the integrand to be zero. Using the chain rule, we have $\frac{dg_{ik}}{d\tau} = \frac{dx^j}{d\tau}\partial_j g_{ik}$. Thus, we require

$$-\frac{1}{2}\partial_k g_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} + \partial_j g_{ik}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} + g_{ik}\frac{d^2x^i}{d\tau^2} = 0. \tag{5.17}$$

Multiplying by the inverse metric g^{lk}, we get our final expression

$$\frac{d^2x^l}{d\tau^2} + \frac{1}{2}g^{lk}\left(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}\right)\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} = 0. \tag{5.18}$$

This is exactly the geodesic equation but with the specific choice of the Christoffel connection

$$\Gamma^l_{ij} = \frac{1}{2}g^{lk}\left(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}\right). \tag{5.19}$$

So, on a manifold the extrema of the proper time (and proper distance) are curves which parallel transport their tangent vector with respect to the Christoffel connection; it doesn't matter if there is another connection defined on the manifold. This is one reason why the Christoffel connection is the one used in General Relativity. It's the connection under which the geodesic equation describes the path of unaccelerated particles.

The geodesic equation can in fact be thought of as a curved space generalisation of Newton's laws, $\vec{f} = m\vec{a}$ for $\vec{f} = 0$. We can even introduce forces on the right hand side in a tensorial way. The equation of motion for a particle of mass m and charge q in GR is

$$\frac{d^2x^i}{d\tau^2} + \Gamma^i_{jk}\frac{dx^j}{d\tau}\frac{dx^k}{d\tau} = \frac{q}{m}F^i{}_l\frac{dx^l}{d\tau}. \tag{5.20}$$

As a last point, we never argued as to why the geodesics maximise the proper time. In Lorentzian spacetime, the character of a geodesic (whether it is time-like, space-like, or null) does not change. If we have a time-like curve, we can approximate it as a series of null curves which have proper time of zero. Timelike geodesics cannot be curves of minimum proper time, since then they would always be infinitesimally close to a null curve as we increase the number of jagged edges.

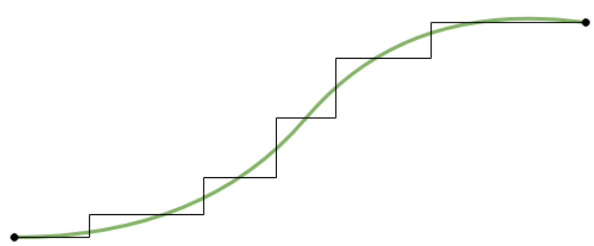

Figure 4:

This is how you can remember which twin in the twin paradox ages more. The one who stays on the Earth's geodesic experiences more proper time.