Trinity College Dublin Senior Sophister Theoretical Physics

Notes on MAU34301 - Differential Geometry

Matthew Dunne

- **Moderators**: Sergey Frolov (Lecturer) and Vera Posch (Tutor)
- **Grade Weighting**: 100% Exam.
- **Textbook References**: Differential Geometry of Manifolds - S. Lovett

Contents

1 Differentiable Manifolds **2**
1.1 Definition of a Manifold . . . 2
1.2 Elements of Topology . . . 4
1.3 Mappings and Tensors on Manifolds . . . 5
1.4 Embeddings and Immersions of Manifolds . . . 10
1.5 Surfaces in Euclidean Space . . . 11

2 Spaces and Groups **14**
2.1 Real Projective Space and Quaternions . . . 14
2.2 Complex Projective Space . . . 16

3 Elements of Lie Algebra **18**
3.1 Algebraic Structure of Lie Groups . . . 18
3.2 Subgroups and Canonical Coordinates . . . 20
3.3 Linear Representations of Groups . . . 22
3.4 Adjoint Representation . . . 23

4 Vector Bundles on a Manifold **26**
4.1 The Tangent Bundle . . . 26

4.2 The Cotangent Bundle . . . 26

5 Vector Fields and Lie Derivatives **27**
5.1 Vector and Tensor Fields . . . 27
5.2 Integral Curves . . . 29
5.3 The Lie Derivative . . . 31
5.4 The Lie Derivative of a Volume Element . . . 33

6 Covariant Differentiation **35**
6.1 The Differential Map on Tensors . . . 35
6.2 New Tensor Unlocked . . . 36
6.3 Parallel Transport of Fields . . . 39

7 Geometry of Connections **42**
7.1 Geodesics . . . 42
7.2 Connections Compatible with the Metric . . . 43
7.3 Geodesics and the Metric . . . 45

8 Curvature Tensors **47**
8.1 Derivation of the Riemann Tensor . . . 47
8.2 Symmetries of The Curvature Tensor . . . 50
8.3 A Few Important Tensors for GR . . . 52

1 Differentiable Manifolds

To begin the course we will reinstate the key results from our mathematical study of manifolds in MAU33206 with Florian Naef. This time we will develop the same ideas in a slightly different way which serves us better as we go on to study advanced topics in theoretical physics.

1.1 Definition of a Manifold

Definition (Region without Boundary / Open Set): A **region** without boundary is an open set D of points in $\mathbb{R}^n$ such that for any point $P_0 \in D$, all points $P_i \in B_\epsilon(P_0)$ are also contained within D for some positive radius ϵ.

Definition (Region with Boundary / Closed Set): For any region without boundary D, the corresponding region *with boundary* is the set D together with all the *boundary points* of D.

Definition ($\mathbb{R}^n$): $\mathbb{R}^n$ or n-dimensional Euclidean space is the set of all points $x = (x^1, \ldots, x^n)$ together with the notion of the *distance* l between any two points being given by

$$l^2(x,y) = \sum_{i=1}^{n}(x^i - y^i)^2. \tag{1.1}$$

Definition (Differentiable n-Manifold): A set M is a differentiable (smooth) manifold of dimension n if it can be written as the union of a finite number of subsets U_q such that;

1. Each U_q has defined on it coordinates x_q^α, $\alpha = 1,\dots,n$ by virtue of which U_q is identifiable (bijective) with a region of $\mathbb{R}^n$. We call each U_q together with its coordinate system a *chart*.

2. Each non-empty intersection $U_q \cap U_p$ between charts is identifiable with regions of $\mathbb{R}^n$ with respect to the restrictions of both (x_q^α) and (x_p^α) to the intersection points and we can use the common domain of both coordinate systems to express the coordinate systems in terms of one another in a one-to-one and differentiable manner (change of variables). Thus if we denote the transition functions between x_p^α and x_q^α as

$$x_p^\alpha = x_p^\alpha(x_q^1,\dots,x_q^n) \qquad \text{and} \qquad x_q^\alpha = x_q^\alpha(x_p^1,\dots,x_p^n) \tag{1.2}$$

then the Jacobian $J_{pq} = \det(\partial x_p^\alpha / \partial x_q^\beta)$ is non-zero on $U_p \cap U_q$.

(Abuse Of) Notation: Here and in what follows we will use x_p^α and x_q^α to denote local coordinates on U_p and U_q, but they will also denote the transition functions between local coordinate systems such that we can effectively write the x_p coordinates as functions of the x_q coordinates and vice versa. This allows us to write derivatives of the transition functions in very convenient way as below

$$\frac{\partial x_p^\alpha}{\partial x_q^\beta} \tag{1.3}$$

where clearly we are using the transition function notation in the numerator and the coordinate notation in the denominator. Using this notation we can immediately write down some useful identities, the first of which is

$$\sum_{\gamma=1}^{\dim M} \frac{\partial x_p^\alpha}{\partial x_q^\gamma}\frac{\partial x_q^\gamma}{\partial x_p^\beta} = \frac{\partial x_p^\alpha}{\partial x_q^\gamma}\frac{\partial x_q^\gamma}{\partial x_p^\beta} = \delta_\beta^\alpha. \tag{1.4}$$

where we use both notations simultaneously and contract the repeated index γ. Another more complicated identity we can write down is

$$\frac{\partial}{\partial x_p^\gamma}\frac{\partial x_p^\alpha}{\partial x_q^\beta} = \frac{\partial^2 x_p^\alpha}{\partial x_p^\gamma \partial x_q^\beta} = \frac{\partial^2 x_p^\alpha}{\partial x_q^\rho \partial x_q^\beta}\frac{\partial x_q^\rho}{\partial x_p^\gamma}. \tag{1.5}$$

Where we note that if the order of differentiation was reversed this would reduce to

$$\frac{\partial^2 x_p^\alpha}{\partial x_p^\gamma \partial x_q^\beta} \neq \frac{\partial^2 x_p^\alpha}{\partial x_q^\beta \partial x_q^\gamma} = \frac{\partial}{\partial x_q^\beta}\frac{\partial x_p^\alpha}{\partial x_p^\gamma} = \frac{\partial}{\partial x_q^\beta}\delta_\gamma^\alpha = 0. \tag{1.6}$$

Now with this language under our control we can start to specify the type of manifolds we are most interested in for the purpose of studying theoretical physics. Namely, we want to be able to specify the algebraic structure that any specific manifold is endowed with. For this we must discuss some aspects of *topology*.

1.2 Elements of Topology

Definition (Topological Space): A topological space is a set X of which certain subsets called **open sets** are distinguished and satisfy:

1. Any finite intersection is again an open set.
2. Any union is again an open set.
3. The trivial sets X and $\varnothing$ are open.

The compliment of an open set is called a closed set. In Euclidean space the topology is that where "*regions*" are open sets.

Definition (Induced Topology): Given any subset $A \in \mathbb{R}^n$, the induced topology on A is that with $A \cap U$ open where U ranges over all open sets of $\mathbb{R}^n$.

Definition (Topologically Equivalent Spaces): Two topological spaces are topologically equivalent or homeomorphic if there is a bijection between them such that both the map and it's inverse are continuous. Such a map is called a *homeomorphism*.

Definition (Topology on a Manifold): The topology on a manifold is given by the following specification of open sets: In every local neighbourhood $U_q \subset M$ *the open regions are said to be open in the topology on* M. The totality of open sets of M is obtained by admitting arbitrary unions of countable collections of such regions to be open.

Definition (Hausdorff Space): A topological space is called *Hausdorff* if any two points are contained in disjoint open sets. Any metric space is Hausdorff because we can always define open balls of radius $\rho(x,y)/3 > 0$ around the points x, y that don't intersect.

Definition (Compact Space): A topological space X is said to be *compact* if every countable collection of open sets covering X has a finite sub-cover also covering X.

Definition (Connected Space): A topological space X is *connected* if any two points on X can be joined by a continuous path.

All of the above definitions are things that we have, in some form or another, come across in analysis modules. Now we can return to our classification of manifolds by defining the orientability of a manifold.

Definition (Oriented Manifold): A manifold M is said to be *oriented* if one can choose its atlas so that for every pair of intersecting neighbourhoods U_q, U_p, the Jacobian

$$J_{pq} = \det\left(\frac{\partial x_p^\alpha}{\partial x_q^\beta}\right) \tag{1.7}$$

of the transition function is *positive*. It follows that *any connected oriented manifold has exactly two orientations, one positive and one negative.*

1.3 Mappings and Tensors on Manifolds

In the following formulation let $M = \cup_p U_p$ be an m-manifold with local coordinates (x_p^α) and $N = \cup_q V_q$ be a n-manifold with local coordinates (y_q^β).

Definition (Smoothly Equivalent Manifolds): The manifolds M and N are said to be smoothly equivalent or *diffeomorphic* if there is a bijection $f : M \to N$ with an inverse $f^{-1} : N \to M$ such that both maps are class C^k for some $k \geq 1$.

Now we will move on to the concept of tangent vectors of a manifold with an intuitive example. Consider the following picture:

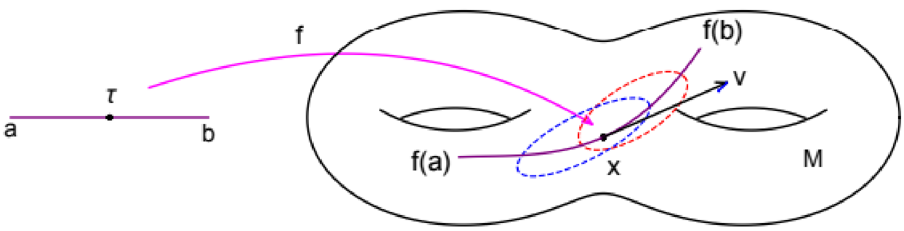


Figure 1.1: A point of a parametric curve on a manifold can be equivalently expressed in either local coordinate system if it lies in the intersection.

In U_p (the blue neighbourhood) we can describe the parametric curve passing through as

$$x_p^\alpha = x_p^\alpha(\tau) \tag{1.8}$$

with the corresponding tangent vector

$$v \equiv \dot{x}_p^\alpha = (\dot{x}_p^1, \ldots, \dot{x}_p^m). \tag{1.9}$$

The same reasoning applies for U_q (the red neighbourhood), and in the intersection we can use transition functions to write the curve in either local coordinate system as a function of the other e.g.

$$x_p^\alpha(x_q^1(\tau),\dots,x_q^m(\tau)) = x_p^\alpha(\tau). \tag{1.10}$$

So using the chain rule we can relate the velocity vectors in in the intersection as

$$\dot{x}_p^\alpha = \frac{\partial x_p^\alpha}{\partial x_q^\beta}\frac{\partial x_q^\beta}{\partial \tau} = \frac{\partial x_p^\alpha}{\partial x_q^\beta}\dot{x}_q^\beta. \tag{1.11}$$

Definition (Tangent Vector to Manifold): A tangent vector to a point $x \in M$ is represented in terms of the local coordinates (x_p^α) by an m-vector ξ^α of components which obey the following transformation law to any other system (x_q^β) of local coordinates

$$\xi_p^\alpha = \left(\frac{\partial x_p^\alpha}{\partial x_q^\beta}\right)_x \xi_q^\beta. \tag{1.12}$$

The set of all tangent vectors to M at x is the *tangent space to* M at x denoted T_xM.

Definition (Push Forward): A smooth map $f : M \to N$ gives rise for each x to an induced linear map known as a *push forward* between tangent spaces which we denote

$$f_* : T_xM \to T_{f(x)}N, \tag{1.13}$$

and define as sending the velocity vector at x of any smooth curve $x(\tau)$ on M to the velocity vector at $f(x)$ of the curve $f(x(\tau))$ on N. So in terms of the local coordinates (x^α) on M and $(y^b) = f^b(x^\alpha)$ on N, the push forward of a tangent vector ξ^α on M is given by

$$\xi^\alpha \mapsto \eta^b = \left(\frac{\partial f^b}{\partial x^\alpha}\right)\xi^\alpha. \tag{1.14}$$

Definition (Riemann Metric): 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 of the points normally denoted $g_{\mu\nu}^{(p)}$ for the local coordinates of the chart U_p. With the metric we define a Riemannian scalar product, $\langle\cdot,\cdot\rangle$, of tangent vectors which is symmetric and coordinate-independent i.e.

$$\langle\xi,\eta\rangle = g_{\alpha\beta}^{(p)}\xi_p^\alpha\eta_p^\beta = \langle\eta,\xi\rangle \quad , \quad \langle\xi,\xi\rangle = |\xi|^2 \quad \text{and} \quad g_{\alpha\beta}^{(p)}\xi_p^\alpha\eta_p^\beta = g_{\alpha\beta}^{(q)}\xi_q^\alpha\eta_q^\beta. \tag{1.15}$$

The coefficients of the metric transform between local coordinates (p) and (q) as

$$g_{\gamma\sigma}^{(q)} = \frac{\partial x_p^\alpha}{\partial x_q^\gamma}\frac{\partial x_p^\beta}{\partial x_q^\sigma}g_{\alpha\beta}^{(p)} \tag{1.16}$$

which we can write in infinitesimal form

$$ds^2 = g_{\alpha\beta}^{(p)}dx_p^\alpha dx_p^\beta = g_{\alpha\beta}^{(q)}dx_q^\alpha dx_q^\beta. \tag{1.17}$$

The quantity ds is called the line element and as we can see it is a chart-invariant distance between two infinitesimally close points.

Definition (Tensors on Manifolds): A tensor of type (k,l) and rank $k+l$ on M is given in each local coordinate system (x_p^i) by a family of functions ${}^{(p)}T^{i_1\dots i_k}_{j_1\dots j_l}(x)$ of the point x. In other local coordinates (x_q^i) the components ${}^{(q)}T^{s_1\dots s_k}_{t_1\dots t_l}(x)$ of the same tensor are given by

$$ {}^{(q)}T^{s_1\dots s_k}_{t_1\dots t_l} = \frac{\partial x_q^{s_1}}{\partial x_p^{i_1}} \cdots \frac{\partial x_q^{s_k}}{\partial x_p^{i_k}} \cdot \frac{\partial x_p^{j_1}}{\partial x_q^{t_1}} \cdots \frac{\partial x_p^{j_l}}{\partial x_q^{t_l}} \cdot {}^{(p)}T^{i_1\dots i_k}_{j_1\dots j_l}. \tag{1.18}$$

Additionally, we define the following operations on and between tensors:

1. **Permutation of indices**: We say that a tensor $\tilde{T}^{i_1\dots i_k}_{j_1\dots j_l}$ is obtained from a tensor $T^{i_1\dots i_k}_{j_1\dots j_l}$ by means of a permutation σ of the lower indices if at each point of M
$$\tilde{T}^{i_1\dots i_k}_{j_1\dots j_l} = T^{i_1\dots i_k}_{\sigma(j_1\dots j_l)}$$
2. **Taking Traces (Index Contraction)**: We can contract a type (k,l) tensor $T^{i_1\dots i_k}_{j_1\dots j_l}$ to a type $(k-1,l-1)$ tensor by $\tilde{T}^{i_1\dots i_{k-1}}_{j_1\dots j_{l-1}}$ choosing to contract the i_a, j_b indices such that
$$\tilde{T}^{i_1\dots i_{k-1}}_{j_1\dots j_{l-1}} = T^{i_1\dots i_{a-1} n \dots i_{a+1}\dots i_k}_{j_1\dots j_{b-1} n j_{b+1}\dots j_l}$$
3. **Tensor Product:** Given two tensors T of type (p,q) and P of type (k,l) we define their product to be the type $(p+k, q+l)$ tensor $S = T \otimes P$ with components given by the equation
$$S^{i_1\dots i_{p+k}}_{j_1\dots j_{q+l}} = T^{i_1\dots i_p}_{j_1\dots j_q} P^{i_{p+1}\dots i_{p+k}}_{j_{q+1}\dots j_{q+l}}.$$

The tensor product *is associative but not commutative.*

Definition (Vectors and Co-Vectors): We refer to tensors ξ^i of type $(1,0)$ as vectors and tensors η_j of type $(0,1)$ as co-vectors or contra-vectors.

Definition (Directional Derivative): We can associate with each vector $\xi = \xi^i$ a linear differential operator which gives the derivative of a function in the direction of the vector ξ given by

$$\partial_\xi = \xi^i \frac{\partial}{\partial x^i}. \tag{1.19}$$

So for an arbitrary function f we can define the scalar quantity

$$\partial_\xi f = \xi^i \frac{\partial f}{\partial x^i} \tag{1.20}$$

to be the *directional derivative* of f in the direction of ξ.

Proposition (Invariance of a Function's Differential): In any coordinate system x^i we can write the differential of a function f as

$$df = \frac{\partial f}{\partial x^i} dx^i. \tag{1.21}$$

Since dx^i is a vector, df has the same value in any coordinate system and this holds in general for any one form being contracted with dx^i. Through this we identify $e^i \equiv dx^i$ with the canonical basis of co-vectors or cotangent space.

Definition (Bilinear Forms): A basis for tensors of type $(0,2)$ at a given point follow naturally as the product $e^i \otimes e^j$. An arbitrary tensor of this type can be regarded as a bilinear form on a pair of vectors since we can write the value of the tensor T_{ij}

$$T_{ij}\xi^i\eta^j \tag{1.22}$$

acting on these vectors as above.

Furthermore we can decompose any tensor of this type into symmetric and alternating contributions using the construction

$$T_{ij} = T_{ij}^{sym} + T_{ij}^{alt}$$

where

$$T_{ij}^{sym} = \frac{1}{2}(T_{ij} + T_j i) = T_{ji}^{sym},$$
$$T_{ij}^{alt} = \frac{1}{2}(T_{ij} - T_{ji}) = -T_{ji}^{alt}.$$

Symmetric and alternating tensors of type $(0,2)$ have the basis tensors

$$\frac{e^i \otimes e^j + e^j \otimes e^i}{2} \equiv dx^i dx^j, \qquad i \leq j$$

and

$$e^i \otimes e^j - e^j \otimes e^i \equiv dx^i \wedge dx^j = -dx^j \wedge dx^i \qquad i < j,$$

respectively.

Definition (Skew-Symmetric (0,k) Tensor): A skew-symmetric tensor of type $(0,k)$ is a tensor $T_{i_1,\ldots,i_k}$ satisfying

$$T_{\sigma(i_1\ldots i_k)} = \mathfrak{s}(\sigma)T_{i_1\ldots i_k} \tag{1.23}$$

where $\mathfrak{s}(\sigma)$ is the sign of the permutation σ. The differential form of the skew-symmetric tensor is

$$T_{i_1\ldots i_k}e^{i_1} \otimes \cdots \otimes e^{i_k} = \sum_{i_1<\cdots<i_k} T_{i_1\ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \frac{1}{k!}T_{i_1\ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}. \tag{1.24}$$

A skew-symmetric tensor of type $(0, k)$ in a k-manifold is completely determined by the component $T_{1,\dots,k}$ since all other components are obtained up to a sign by permuting the indices. We define the skew-symmetric tensor of type $(0, n)$ given by $T_{1\dots n} = 1$ with the notation

$$\epsilon_{i_1\dots i_n}. \tag{1.25}$$

It is called the *Levi-Civita symbol* of rank n. We use the word symbol because this tensor holds its definition in all charts, while a general tensor would only be defined as such in some charts.

Theorem (Transformation of (0,n) Skew Tensors): Skew-symmetric tensors of type $(0, n)$ where n is the dimension of the manifold M transform between charts as

$$^{(p)}T_{1\dots n} =^{(q)} T_{1\dots n} \cdot J \tag{1.26}$$

where J is the Jacobian $\det(\partial x^i_q / \partial x^j_p)$.

Definition (Manifold Metric): A metric g_{ij} on a manifold is a tensor of type $(0, 2)$, and on an oriented manifold M with dimension m such a metric gives an expression for the volume element

$$T_{i_1,\dots,i_n} = \sqrt{|g|}\epsilon_{i_1\dots i_m} \qquad \text{where} \qquad g = \det(g_{ij}). \tag{1.27}$$

We can also write this element in differential notation

$$\Omega = \sqrt{|g|} = dx^1 \wedge \cdots \wedge dx^m. \tag{1.28}$$

In the case the g_{ij} is Riemannian the volume V of M is

$$V = \int_M \Omega = \int_M \sqrt{g} dx^1 \wedge \cdots \wedge dx^m. \tag{1.29}$$

Definition (Distance Function): A Riemann metric $ds^2 = g_{ij}dx^i dx^j$ on a connected manifold M with a distance function $\rho(x, y)$ defined by the path of least distance

$$\rho(x, y) = \inf_\gamma \int_\gamma ds. \tag{1.30}$$

In other words this infimum is taken over all piece-wise smooth curves joining the points x and y. The topology on M defined by this metric space structure coincides with the Euclidean topology on M.

Definition (Push-Forward): A smooth map $f : M \to N$ gives rise for all points $x \in M$ to the push-forward map of tangent spaces

$$f_* : T_x M \to T_{f(x)} N \tag{1.31}$$

which in terms of local coordinates x^i of a chart U containing x and y^j of a chart V containing y such that $y^j = f^j(x^1, \ldots, x^m)$ is written as

$$f_* : \xi^i \to \eta^j = \frac{\partial f^j}{\partial x^i}\xi^i. \tag{1.32}$$

And for any other tensors of type $(k, 0)$ this generalises to

$$f_* : \xi^{i_1 \ldots i_k} \to \eta^{j_1 \ldots j_k} = \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_k}}{\partial x^{i_k}} \xi^{i_1 \ldots i_k}. \tag{1.33}$$

The analogue for tensors of type $(0, k)$ is the *pullback* which we are already very familiar with from Calculus on Manifolds.

Remark (Push-Forward and Pull-back): If we have both a type $(0, k)$ tensor $\eta_{j_1 \ldots j_k}$ and a type $(k, 0)$ tensor $\xi^{i_1 \ldots i_k}$ in $T^{(0,k)}_{f(x)}N$ and $T^{(k,0)}_x M$ respectively, then

$$\begin{aligned}(f^*\eta)(\xi) &= \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_k}}{\partial x^{i_k}} \eta_{j_1 \ldots j_k} \xi^{i_1 \ldots i_k} \\ &= \eta_{j_1 \ldots j_k} \frac{\partial f^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial f^{j_k}}{\partial x^{i_k}} \xi^{i_1 \ldots i_k} = \eta(f_* \xi).\end{aligned} \tag{1.34}$$

1.4 Embeddings and Immersions of Manifolds

Definition (Immersions): A manifold M is said to be immersed in a manifold N of dimension $n \geq m$ if there is a smooth map $f : M \to N$ such that the push-forward map f_* is at each point a one-to-one map of the tangent space. In this case, the map f is called an *immersion* of M into N. In terms of local coordinates, this means the Jacobian matrix of f has rank m at each point of M.

Definition (Embedding): An immersion of M into N is called an embedding if f itself is one-to-one. Then M is called a submanifold of N.

Definition (Induced Metric on a Submanifold): Let M be a submanifold of the Riemann manifold N with metric $g^{(N)}_{ij}$, then the pull-back of this metric to M yields the tensor

$$g^{(M)}_{kl}(x) = \frac{\partial f^i}{\partial x^k}\frac{\partial f^j}{\partial x^l} g^{(N)}_{kl}(f(x)), \qquad k, l = 1, \ldots, m \qquad i, j = 1, \ldots, n \tag{1.35}$$

which is called the *induced metric* on the submanifold M by the metric of N.

We can do something similar for the line element of N which we will remind ourselves is $ds_N^2 = g^{(N)}_{ij}(y)dy^i dy^j$. Now letting $y^i = f^i(x^1, \ldots, x^m)$. Then the line element can be transformed as

$$ds_N^2 = g^{(N)}_{ij}(f(x))\frac{\partial f^i}{\partial x^k}dx^k \frac{\partial f^j}{\partial x^l}dx^l = g^{(M)}_{kl}(x)dx^k dx^l = ds_M^2. \tag{1.36}$$

Thus the infinitesimal distances measured by each metric on their respective manifold are the same.

Theorem (Pullback of a Skew-Symm Form to a Submanifold): The pullback of the skew-symmetric form

$$\frac{1}{m!}T_{a_1\dots a_m}dy^{a_1}\wedge\dots\wedge dy^{a_m}$$

to the m-dimensional submanifold with charts related as $y^a = f^a(x^1,\dots,x^m)$ is given by

$$\left(\frac{1}{m!}J^{a_1\dots a_m}T_{a_1\dots a_m}\right)dx^1\wedge\dots\wedge dx^m \tag{1.37}$$

where $J^{a_1,\dots,a_m}$ is the $m\times m$ minor of the matrix $(\partial y^a/\partial x^i)$ formed from the columns numbered $a_1,\dots,a_m$.

Using Local Coordinates on a Submanifold: We shall always assume that any submanifold M is defined in each chart U_p of the containing manifold N by a system of equations

$$f_p^i(x_p^1,\dots,x_p^n) = 0, \qquad \text{rank}\left(\frac{\partial f_p^i}{\partial x_p^\alpha}\right) = n-m \implies i = 1,\dots,n-m \tag{1.38}$$

with the property that on each intersection $U_p\cap U_q$ the systems $(f_p^i = 0)$ and $(f_q^i = 0)$ have the same set of zeroes. We can always introduced coordinates $y_p^1,\dots,y_p^n$ on $U_p\subset N$ satisfying

$$y_p^{m+1} = f_p^1(x_p^1,\dots,x_p^n) = 0, \quad y_p^{m+2} = f_p^2(x_p^1,\dots,x_p^n), \quad y_p^n = f_p^{n-m}(x_p^1,\dots,x_p^n).$$

Then all coordinates corresponding to a dimension higher than m vanish and the $1\to m$ coordinates y_p^α serve as local coordinates on M.

Definition (Manifold with Boundary): A closed region A of a manifold M defined by an inequality

$$f(x)\leq 0 \qquad (\text{or } f(x)\geq 0), \tag{1.39}$$

where f is a real-valued function on M is called a *manifold with boundary*. It is assumed that the boundary $\partial A = \{x\in A \mid f(x) = 0\}$ is a non-singular submanifold of M.

Definition (Closed Manifold): A compact manifold without boundary is called *closed*.

1.5 Surfaces in Euclidean Space

Definition (Non-Singular Surface): A *non-singular surface* M in $\mathbb{R}^n$ is given by a set of $n-m$ equations

$$f^i(x^1,\dots,x^n) = 0 \qquad i = 1,\dots,n-m$$

where for any point x the matrix $(\partial f_i/\partial x^\alpha)$ has rank $n - m$.

(NB) Theorem (Valid Local Coordinates for a Non-Singular Surface in $\mathbb{R}^n$): Take the above definition, now let $J_{j_1 \cdots j_{n-m}}$ be the minor of a sub-matrix made up of the columns of $(\partial f_i/\partial x^\alpha)$ which are indexed by $j_1, \ldots, j_{n-m}$. Let $U_{j_1 \ldots j_{n-m}}$ be the regions consisting of all points of the surface at which $J_{j_1 \cdots j_{n-m}}$ is nonzero. Since the surface is non-singular, it must then be covered by the regions $U_{j_1 \cdots j_{n-m}}$ such that

$$M = \bigcup_{j_1 \cdots j_{n-m}} U_{j_1 \cdots j_{n-m}}.$$

Let $x_0 \in M$, if at x_0 the minor is nonzero, then we can take local coordinates around x_0 to be

$$(y^1, \ldots, y^m) = (x^1, \ldots, \hat{x}^{j_1}, \ldots, \hat{x}^{j_{n-m}}, x^{n-m+1}, \ldots, x^n). \tag{1.40}$$

where we omit the hatted coordinates. (Prove it for exam???).

Theorem (Above Coordinates (1.40) **make M a Smooth Manifold):** The covering of the non-singular surface M defined as in the above theorem by the regions

$$U_{j_1 \cdots j_{n-m}}, \qquad 1 \le j_1 < \cdots < j_{n-m} \le n,$$

each furnished with local coordinates (1.40) defines on the surface the structure of a smooth manifold.

Proof: Since the Jacobian of the minor is full rank we can solve the equations $f_i = 0$ for x^{j_i}.

$$\begin{aligned} (y^1, \ldots, y^m) &= (x^1, \ldots, \hat{x}^{j_1}, \ldots, \hat{x}^{j_{n-m}}, \ldots x^n) \\ \implies x^{j_i} &= \varphi^i(y^1, \ldots, y^m), \qquad i = 1, \ldots, n-m. \end{aligned}$$

Similarly in the chart $U_{s_1 \ldots s_{n-m}}$ with coordinates

$$\begin{aligned} (z^1, \ldots, z^m) &= (x^1, \ldots, \hat{x}^{s_1}, \ldots, \hat{x}^{s_{n-m}}, \ldots x^n) \\ \implies x^{s_i} &= \psi^i(y^1, \ldots, y^m), \qquad i = 1, \ldots, n-m \end{aligned}$$

where φ^i and ψ^i are both smooth functions. Now to show that the transition functions are diffeomorphisms in $U_{j_1 \cdots j_{n-m}} \cap U_{s_1 \ldots s_{n-m}}$ (for simplicity we will assume $1 < j_1 < s_1 < j_2 < \cdots$). We can observe how each component in each set of coordinates are related to each other via tabulation:

$$\begin{array}{cc}
x^1 = & y^1 = z^1 \\
\cdots\cdots\cdots & \cdots\cdots\cdots \\
x^{j_1-1} = & y^{j_1-1} = z^{j_1-1} \\
x^{j_1} = & \varphi^1(y^1, \ldots, y^m) = z^{j_1} \\
x^{j_1+1} = & y^{j_1} = z^{j_1+1} \\
\cdots\cdots\cdots & \cdots\cdots\cdots \\
x^{s_1-1} = & y^{s_1-2} = z^{s_1-1} \\
x^{s_1} = & y^{s_1-1} = \psi^1(z^1, \ldots, z^m) \\
x^{s_1+1} = & y^{s_1} = z^{s_1} \\
\cdots\cdots\cdots & \cdots\cdots\cdots \\
x^m = & y^m = z^m
\end{array}$$

Which shows that the two transition functions are mutual inverses as needed.

Remarks: The Jacobian of the transition function $y \to z$ is

$$J_{(y)\to(z)} = \pm \frac{J_{s_1 \dots s_{n-m}}}{J_{j_1 \dots j_{n-m}}}.$$

The tangent space to the surface M is identifiable with the linear subspace of $\mathbb{R}^n$ consisting of solutions of the equations

$$\frac{\partial f_1}{\partial x^\alpha} \xi^\alpha = 0, \dots, \frac{\partial f_{n-m}}{\partial x^\alpha} \xi^\alpha = 0.$$

Thus the one forms ∇f_i are orthogonal to the surface at each point.

Definition (Orientation Class): For a surface M, there are ordered bases $\tau = (e_1, \dots, e_m)$ for the tangent space to M at any point x. Any two such bases are related by a (full rank) linear transformation A. We say the ordered bases, τ_1, τ_2 lie in the *same orientation class* if $\det A > 0$ or the *opposite orientation class* if $\det A < 0$.

Definition (Alternate form of Orientability): A manifold is *orientable* if it is possible at every point to choose a single orientation class depending continuously on the points. Such a choice is called an orientation.

This definition and the previous definition are equivalent and it is easy to see that they imply each other. For example in the previous definition we choose a set of standard basis vectors tangent to the local coordinate axes. For any overlapping chart, this orienting frame will be related to the other coordinates by the Jacobian which is always positive in the intersection of charts on a manifold.

Theorem (Orientable Surface): A smooth non-singular surface M in $\mathbb{R}^n$ with $m < n$ defined by a system of equations is orientable.

Definition (Two Sided Manifold): A connected submanifold of $\mathbb{R}^n$ of dimension $(n-1)$ is called *two sided* if a single-valued continuous field of unit normals can be defined on it.

Theorem (Two-sided $\implies$ Orientable): Any two-sided hypersurface in $\mathbb{R}^n$ is orientable.

Proof: It can be shown that any closed hypersurface in $\mathbb{R}^n$ is defined by a single non-singular equation of the form $f(x) = 0$. Thus it bounds a manifold with boundary and one can prove that any closed hypersurface is two-sided.

Now we want to proceed to talk about how surfaces relate to transformation groups since this topic will be very useful in physics. The definition of a group is as we have seen before.

Definition (Group): A group G is a non-empty set on which there is a defined binary operation $(a,b) \to ab$ satisfying the following *group axioms*

1. **Closure:** If $a, b \in G$ then $ab \in G$.
2. **Associativity:** If $a, b, c \in G$, then $a(bc) = (ab)c$.
3. **Identity:** There is an element $\mathbf{1} \in G$ such that $a\mathbf{1} = \mathbf{1}a = a$ for all $a \in G$.
4. **Inverse:** If $a \in G$ then there exists an element $a^{-1} \in G$ such that $aa^{-1} = a^{-1}a = \mathbf{1}$.

Definition (Lie Group): A manifold G is called a Lie group if it has given on it a group operation with the property that the two maps $\varphi : G \to G$, defined by $\varphi(g) = g^{-1}$, $\psi : (G \times G) \to G$ defined by $\psi(g,h) = gh$ are smooth.

2 Spaces and Groups

2.1 Real Projective Space and Quaternions

Definition ($\mathbb{R}P^n$): The real projective space $\mathbb{R}P^n$ is the set of all straight lines in $\mathbb{R}^{n+1}$ passing through the origin. It can also be referred to as the set of all equivalence classes of non-zero vectors in $\mathbb{R}^{n+1}$ (vectors that are the same modulo a rescaling).

Each line passing through the origin obviously passes through two diametrically opposite points on the sphere S^n and so $\mathbb{R}P^n$ corresponds one to one with pairs of diametrically opposed points on $\mathbb{R}^n$. Thus $\mathbb{R}P^n \cong S^n/Z_2$.

Definition (Quaternions): The set $\mathbb{H}$ of *quaternions* consists of all linear combinations

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}, \qquad a, b, c, d \in \mathbb{R} \tag{2.1}$$

where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are all linearly independent vectors that square to -1. We define the multiplication of the basis elements as

$$\mathbf{ij} = \mathbf{k}, \qquad \mathbf{jk} = \mathbf{i}, \qquad \mathbf{ki} = \mathbf{j}.$$

with the property of anti-commutativity. Then $\mathbb{H}$ is an associative algebra over the field of real numbers. For each quaternion we can define a matrix whose determinant is the length of the vector as

$$A(q) = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}, \qquad A(q) \in \mathrm{Mat}(2, \mathbb{C}). \tag{2.2}$$

We also define the operation of conjugation on $\mathbb{H}$ as

$$\bar{q} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}. \tag{2.3}$$

Quaternion Lemma 1: The map $q \mapsto A(q)$ is one to one and

$$A(q_1 q_2) = A(q_1)A(q_2)$$

which makes the map an algebra monomorphism (injective homomorphism). We note that the codomain of the map can be expanded in a basis of Pauli matrices since

$$A(\mathbf{i}) = i\sigma_3, \qquad A(\mathbf{j}) = i\sigma_2 \qquad A(\mathbf{k}) = i\sigma_1.$$

Quaternion Lemma 2: The map $q \mapsto \bar{q}$ is an anti-isomorphism of $\mathbb{H}$ meaning it is linear and $\overline{q_1 q_2} = \bar{q}_1 \bar{q}_2$. We can then define the *norm* of a quaternion by

$$\det A(q) = |q|^2 = q\bar{q} = a^2 + b^2 + c^2 + d^2 \geq 0, \tag{2.4}$$

it follows that there is a multiplicative inverse for every nonzero quaternion given by

$$q^{-1} = \frac{\bar{q}}{|q|^2}.$$

Quaternion Lemma 3: The algebra $\mathbb{H}$ of quaternions is a *division algebra* meaning that each non-zero quaternion satisfies

$$qq^{-1} = \mathbf{1} = q^{-1}q,$$

we also have that

$$|q|^2 = \det A(q) \implies |q_1 q_2|^2 = |q_1|^2 |q_2|^2.$$

Hence the set of quaternions with unit norm forms a group under multiplication denoted $\mathbb{H}_1$ in which the elements are such that $q^{-1} = \bar{q}$. Noting that $\mathbb{H}$ can be identified with $\mathbb{R}^4$ we can think of $\mathbb{H}_1$ as the hypersurface formed by points at a distance of 1 from the origin which is clearly the definition of $S^3 \cong SU(2)$.

Quaternion Lemma 4: If $|q|^2 = 1$, then the transformation defined by

$$\alpha_q : x \mapsto qxq^{-1}, x \in \mathbb{H}_0 = \mathbb{R}^3$$

is a rotation of $\mathbb{R}^3$. In other words the map $q \mapsto \alpha_q$ is a homomorphism from $\mathbb{H}_1 \cong SU(2)$ to the group of rotations of $\mathbb{R}^3$. Thus $SO(3)$ is isomorphic to $SU(2)/Z_2 \cong S^3/Z_2 \cong \mathbb{R}P^3$.

Definition (Manifold Structure on $\mathbb{R}P^n$): $\mathbb{R}P^n$ is the set of equivalence classes of nonzero vectors in $\mathbb{R}^{n+1}$ with coordinates $y^0, \ldots, y^n$. For each (not quaternion) $q = 0, 1, \ldots, n$ let U_q denote the set of equivalence classes of vectors $(y^0, \ldots, y^n)$ with $y^q \neq 0$, then clearly $\mathbb{R}P^n = \cup_{q=0}^n U_q$. We introduce the following local coordinates on U_0

$$x_0^i = \frac{y^i}{y^0}, \qquad i = 1, \ldots, n.$$

Now we obtain the local coordinates on each U_q through the recursion relation

$$x^i_{q+1} = \frac{x^{i+1}_q}{x^1_q}, \qquad i = 1, \ldots, n-1, \qquad x^n_{q+1} = -\frac{1}{x^1_q} \tag{2.5}$$

from which we obtain

$$\begin{aligned} x^1_q = \frac{x^2_{q-1}}{x^1_{q-1}} = \frac{x^3_{q-2}x^1_{q-2}}{x^2_{q-2}x^1_{q-2}} = \frac{x^3_{q-2}}{x^2_{q-2}} = \frac{x^4_{q-3}x^1_{q-3}}{x^3_{q-3}x^1_{q-3}} = \frac{x^4_{q-3}}{x^3_{q-3}} = \cdots \\ = \frac{x^{p+1}_{q-p}}{x^p_{q-p}} = \frac{x^{q+1}_0}{x^q_0} = \frac{y^{q+1}}{y^q}. \end{aligned}$$

Therefore $x^1_q \neq 0$ on U_{q+1} and up to a sign x^i_{q+1} can be expressed as the ratio y^{n_i}/y^{q+1} where n_i depends on both i and q. Equations (2.5) also provide us with transition functions on the intersection $U_q \cap U_{q+1}$ for which the Jacobian is

$$J_{(x_q)\to(x_{q+1})} = \left(-\frac{1}{x^1_q}\right)^{n+1} \neq 0.$$

The Jacobian for arbitrary overlaps also appears in recurring form due to the nature of the charts such that

$$\begin{aligned} J_{(x_q)\to(x_p)} &= J_{(x_q)\to(x_{q+1})}J_{(x_{q+1})\to(x_{q+2})}\cdots J_{(x_{p-1})\to(x_p)} \\ = \prod_{k=q}^{p-1}\left(-\frac{1}{x^1_k}\right)^{n+1} &= \prod_{k=q}^{p-1}\left(-\frac{y^k}{y^{k+1}}\right)^{n+1} = \left((-1)^{p-q}\frac{y^q}{y^p}\right)^{n+1} \neq 0. \end{aligned}$$

This is also clearly a smooth function on $U_q \cap U_p$ so we conclude that the real projective space is a smooth manifold. As an aside, it is oriented for odd n.

2.2 Complex Projective Space

Definition (Complex Projective Space): The *complex projective space* $\mathbb{C}P^n$ is the set of equivalence classes of nonzero vectors in $\mathbb{C}^{n+1}$ where two nonzero vectors are equivalent if they are scalar multiples of one another. The charts are defined as in the real case, making $\mathbb{C}P^n$ a $2n$-dimensional smooth manifold.

Definition (Complex Projective Line): The complex projective line $\mathbb{C}P^1$ is the set of points that are equivalence classes of nonzero pairs

$$(z^0, z^1) \propto (\lambda z^0, \lambda z^1), \qquad \lambda \in \mathbb{C} \neq 0.$$

Now consider the function

$$w_0(z^0, z^1) = \frac{z^1}{z^0}$$

defined on $U_0 = \mathbb{C}P^1/\{(0,1)\}$. This covers all points in $\mathbb{C}P^1$ except the equivalence class of $(0,1)$. At this excluded point we define $w_0 = \infty$. Then through this function $\mathbb{C}P^1$ becomes identified with the *extended complex plane*.

Theorem ($\mathbb{C}P^1 \cong S^2$): The complex projective line is diffeomorphic to the sphere S^2.

Proof: On $U_0 : z^0 \neq 0$ we introduce local coordinates u_0, v_0 as

$$u_0 + iv_0 = w_0 = z^1/z^0$$

which define a one to one map onto $\mathbb{R}^2$. On $U_1 : z^1 \neq 0$ we introduce local coordinates u_1, v_1 as

$$u_1 + iv_1 = w_1 = z^0/z^1.$$

Clearly, $\mathbb{C}P^1 = U_0 \cup U_1$. The transition functions in the intersection are

$$w_1 = \frac{1}{w_0} = \frac{u_0 - iv_0}{u_0^2 + v_0^2} \implies (u_1, v_1) = \left(\frac{u_0}{u_0^2 + v_0^2}, -\frac{v_0}{u_0^2 + v_0^2}\right).$$

These are the same functions we used for S^2 under the stereographic projections onto the plane $z = 0$ and thus $\mathbb{C}P^1 \cong S^2$. Through this result we often call the extended complex plane the Riemann sphere. On this sphere $u + iv$ provide local coordinates on the finite part and $1/w$ provide local coordinates for a neighbourhood of the point at infinity.

Definition (The General complex projective space): From each equivalence class of $(n+1)$-dimensional vectors $(z^0, z^1, \ldots, z^n)$ we choose a representative that has unit norm. Then we can form the equation

$$|z^0|^2 + |z^1|^2 + \cdots + |z^n|^2 = 1$$

which defines the unit sphere S^{2n+1} in $\mathbb{C}^{n+1} \cong \mathbb{R}^{2n+2}$. This representation is not unique however since we can multiply each z^i by a pure phase from $U(1)$.

The complex projective space $\mathbb{C}P^n$ can be obtained from the unit sphere

$$S^{2n+1} = \{z \mid \sum_{i=0}^{n} |z^i|^2 = 1\}$$

by identifying all points differing by a scalar factor of the form $e^{i\varphi}$, this identification provides a map

$$S^{2n+1} \mapsto \mathbb{C}P^n \cong S^{2n+1}/U(1),$$

such that the preimage of each point of $\mathbb{C}P^n$ is topologically equivalent to S^1. In particular, since $\mathbb{C}P^1 \cong S^2$ we get a map

$$S^3 \mapsto S^2, \qquad (z^0, z^1) \mapsto w = \frac{z^1}{z^0}, \qquad |z^0|^2 + |z^1|^2 = 1.$$

This map is called the Hopf bundle or the *fibration map*.

3 Elements of Lie Algebra

3.1 Algebraic Structure of Lie Groups

Recall the definition of a Lie group, we will now develop a proper formalism for dealing with Lie groups and the manifold on which they are defined. Let G be a Lie group with identity element $\mathbf{1} = g_0 = (0, \dots, 0)$ and let $T = T_{(1)}$ be the tangent space at this point. We can express the group operations on G in a chart $g_0 \in U_0$ in terms of local coordinates on the manifold as follows:

We choose the coordinates in U_0 so that the identity element is at the origin $\mathbf{1} = g_0 = (0, \dots, 0)$. Now let

$$g_1 = (x^1, \dots, x^n), \qquad g_2 = (y^1, \dots, y^n), \qquad g_3 = (z^1, \dots, z^n)$$

such that all products of these points under the group multiplication are closed in U_0. Then define the product of two elements as

$$g_1 g_2 = (\psi^1(x,y), \psi^2(x,y), \dots, \psi^n(x,y)) = (\psi^i(x,y))$$

where

$$\psi^i(x,y) = \psi^i(x^1, \dots, x^n, y^1, \dots, y^n), \qquad i = 1, \dots, n.$$

The inverse elements will be defined as

$$g_1^{-1} = (\varphi^1(x), \varphi^2(x), \dots, \varphi^n(x)) = (\varphi^i(x)),$$

where

$$\varphi^i(x) = \varphi^i(x^1, \dots, x^n), \qquad i = 1, \dots, n.$$

From each of the group axioms we can then derive an identity starting with the identity, we have that

$$\psi^i(x,0) = \psi^i(0,x) = x^i. \tag{3.1}$$

Next we can use that each element has an inverse to write

$$\psi^i(x, \varphi(x)) = 0. \tag{3.2}$$

Lastly and most importantly we can use associativity to write that $g_1(g_2 g_3) = (g_1 g_2) g_3$ or in coordinate form

$$\psi^i(x, \psi(y,z)) = \psi^i(\psi(x,y), z). \tag{3.3}$$

These results will be very useful in deriving the following few results.

Lemma: Let $\psi(x,y)$ be sufficiently smooth such that it can be written as a series expansion. Then

$$\psi^i(x,y) = x^i + y^i + b^i_{jk} x^j y^k + \text{terms of order} \geq 3$$

and

$$b^i_{jk} = \frac{\partial^2\psi^i}{\partial x^j \partial y^k}\bigg|_{x=y=0}$$

Proof: The Taylor series expansion is

$$\psi^i(x,y) = \psi^i(0,0) + \frac{\partial\psi^i}{\partial x^j}\bigg|_{x=y=0} x^j + \frac{\partial\psi^i}{\partial y^j}\bigg|_{x=y=0} y^j$$

$$+\frac{1}{2}\frac{\partial^2\psi}{\partial x^j\partial x^k}x^j x^k + \frac{1}{2}\frac{\partial^2\psi^i}{\partial y^j\partial y^k}\bigg|_{x=y=0} y^j y^k + \frac{\partial^2}{\partial x^j\partial y^k}\bigg|_{x=y=0} + \cdots.$$

From the first identity above one finds that

$$\psi^i(0,0) = 0, \qquad \frac{\partial\psi^i}{\partial x^j}\bigg|_{x=y=0} = \frac{\partial\psi^i}{\partial y^j}\bigg|_{x=y=0} = \delta^i_j, \qquad \frac{\partial^2\psi^i}{\partial x^j\partial x^k}\bigg|_{x=y=0} = \frac{\partial^2}{\partial y^j\partial y^k}\bigg|_{x=y=0} = 0.$$

Let $\xi, \eta \in T$ with components in terms of x^i denoted by ξ^i and η^i. Then we define their commutator as follows

Definition (Commutator): The *commutator* $[\xi,\eta] \in T$ is defined by

$$[\xi,\eta]^i = c^i_{jk}\xi^j\eta^k, \qquad c^i_{jk} \equiv b^i_{jk} - b^i_{kj} \tag{3.4}$$

and it satisfies the properties:

1. $[,]$ is **bilinear** on the n-dim vector space T.
2. **Skew-symmetry:** $[\xi,\eta] = -[\eta,\xi]$.
3. **Jacobi Identity:** $[[\xi,\eta],\zeta] + [[\zeta,\xi],\eta] + [[\eta,\zeta],\xi] = 0$.

The proofs of the first two properties are trivial but Jacobi's identity has a highly non-trivial structure and its proof is enlightening.

Proof of JI: We start by explicitly writing out both sides of the associativity equation (3.2).

$$\psi^i(x,\psi(y,z)) = x^i + \psi^i(y,z) + b^i_{jk}x^j\psi^k(y,z) + \cdots,$$

$$= x^i + y^i + z^i + b^i_{jk}y^j z^k + b^i_{jk}x^j y^k + b^i_{jk}x^j z^k + b^i_{jk}b^k_{mn}x^j y^m z^n + \cdots$$

On the other hand we have

$$\psi^i(\psi(x,y),z) = \psi^i(x,y) + z^i + b^i_{jk}\psi^j(x,y)z^k + \cdots$$

$$= x^i + y^i + b^i_{jk}x^j y^k + z^i + b^i_{jk}x^j z^k + b^i_{jk}y^j z^k + b^i_{jk}b^k_{mn}x^m y^n z^k + \cdots$$

Cancelling on both sides leaves

$$b^i_{jk}b^k_{mn}x^j y^m z^n - b^i_{jk}b^j_{mn}x^m y^n z^k = (b^i_{jp}b^p_{mn} - b^i_{pn}b^p_{jm})x^j y^m z^n = 0,$$

and hence

$$b^i_{jp}b^p_{mn} - b^i_{pn}b^p_{jm} = 0. \tag{3.5}$$

Computing the Jacobi identity explicitly reveals that the terms reduce to this exact structure and hence the identity holds.

Definition (Lie Algebra): A *Lie algebra* is a vector space $\mathcal{G}$ over a field $\mathbb{K}$ with a bilinear operation $[\cdot,\cdot] : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ called a Lie bracket or commutator such that the operation is skew symmetric and satisfies the Jacobi identity. The *tangent space* of a Lie group G at the identity is with respect to the commutator the corresponding Lie algebra of the Lie group G.

Now lets discuss the tangent space mentioned above in more detail. Let

$$e_1 = \frac{\partial}{\partial x^1}, \dots, e_n = \frac{\partial}{\partial x^n}$$

be the standard basis vectors of T in terms of the coordinates $x^1, \dots, x^n$. We can contract components of the commutator along these basis vectors as

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

If we choose $\xi = e_j,\ \eta = e_k$, then combined with the fact that $(e_m)^n = \delta^n_m$, then we get

$$[e_j, e_k] = c^i_{jk}e_i. \tag{3.6}$$

Definition (Structure Constants): The constants c^i_{jk} which determine the commutation operation on a Lie algebra and are anti-symmetric in their indices are called the *structure constants* of the Lie algebra.

3.2 Subgroups and Canonical Coordinates

Definition (One Parameter Subgroup): A one parameter subgroup of a Lie Group G is defined to be a parametric curve $F(t)$ on the manifold G such that

$$F(0) = 1, \qquad F(t_1 + t_2) = F(t_1)F(t_2), \qquad F(-t) = F(t)^{-1}. \tag{3.7}$$

The velocity vector at the point $F(t)$ on the subgroup is defined by the limiting equation

$$\frac{dF}{dt} = \left.\frac{dF(t+\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \frac{dF}{d\epsilon}(F(t)F(\epsilon))|_{\epsilon=0} = F(t)\left.\frac{dF(\epsilon)}{d\epsilon}\right|_{\epsilon=0}.$$

So $\dot{F}(t) = F(t)\dot{F}(0)$ from which it is clear that the induced action of left multiplication by $F(t)^{-1}$ sends the tangent vector $\dot{F}(t)$ to $\dot{F}(0) = \text{constant} \in T$. Conversely this means that for any element of the tangent space at the identity, $A \in T$ the equation

$$F(t)^{-1}\dot{F}(t) = A$$

is satisfied by a unique one parameter subgroup $F(t)$ of G. If G is a matrix group then the particular form of the function F will always be

$$F(t) = \exp(At). \tag{3.8}$$

Let's discuss now how we can discuss a push forward map can be constructed for the one parameter subgroup. Let $F(t) \in U_0$ have local coordinates

$$(f^1(t), \dots, f^n(t)).$$

Using the notation of group elements in terms of local coordinates, these component functions of the one parameter subgroup will satisfy

$$f^i(0) = 1, \qquad f^i(t_1 + t_2) = \psi^i(f(t_1), f(t_2)), \qquad f^i(-t) = \varphi^i(f(t)).$$

The action on group manifold elements of left multiplication by $F(t)$ is given by

$$x \mapsto y : \quad y^i = \psi^i(f(t), x)$$

This has an induced push forward map given by

$$F_*(t) : \quad \xi^i \mapsto \eta^i = \frac{\partial \psi^i(f(t), x)}{\partial x^j} \xi^j, \qquad \xi \in T_x G, \quad \eta \in T_y G.$$

In terms of local coordinates we can now write the velocity vector as $\dot{F}(t) = (\dot{f}^1(t), \dots, \dot{f}^n(t),$ where

$$\dot{f}^i(t) = \left.\frac{df^i(t+\epsilon)}{d\epsilon}\right|_{\epsilon=0} = \left.\frac{d\psi^i(f(t), f(\epsilon))}{d\epsilon}\right|_{\epsilon=0} = \left.\frac{\partial \psi^i(f(t), x)}{\partial x^j}\right|_{x=0} \cdot \dot{f}^j(0).$$

We can pick out the inverse of the push forward map by conducting this analysis in the opposite direction

$$\dot{f}^i(0) = \left.\frac{df^i(-t+\epsilon)}{d\epsilon}\right|_{\epsilon=t} = \left.\frac{d\psi^i(f(-t), f(\epsilon))}{d\epsilon}\right|_{\epsilon=t} = \left.\frac{\partial \psi^i(\varphi(f(t)), x)}{\partial x^j}\right|_{x=f(t)} \cdot \dot{f}^j(t).$$

Therefore $F_*(t)^{-1}$ sends $\dot{F}(t) \to \dot{F}(0)$, and $F_*(t)$ sends $\dot{F}(0) \to \dot{F}(t)$.

We study one parameter subgroups because they can be used to define *canonical coordinates* in a neighbourhood of the identity of a Lie Group G.

Definition (Canonical Coordinates): Let $A_1, \dots, A_n$ form a basis for the Lie algebra T then for all $A = \sum_i A_i x^i \in T$ there exists a one parameter group defined by the equation $F(t) = \exp(At)$. We set $F(1) = \exp A$ with the summation coefficients $(x^1, \dots, x^n)$ giving us a system of local coordinates in a sufficiently small neighbourhood of the identity $g_0 = 1 \in G$. This choice of coordinates is called the *canonical coordinates of the first kind.*

Another coordinate system can be obtained by choosing $F_i(t) = \exp A_i t$ such that a point in the neighbourhood of the identity is represented as

$$g = F_1(t_1)F_2(t_2)\cdots F_n(t_n)$$

for small $t_1, \ldots, t_n$. Assigning coordinates $x^1 = t_1, \; \cdots x^n = t_n$ we get the *canonical coordinates of the second kind.*

Theorem (Determination of Group Multiplication from Algebra Structure): If the functions $\psi^i(x, y)$ defining the multiplication of points x, y of a Lie group G are real analytic, then in some neighbourhood of g_0 the structure of the Lie algebra determines the multiplication in G.

3.3 Linear Representations of Groups

Definition (Linear Representation): A linear representation of a group G of dimension n is a homomorphism

$$\rho : G \mapsto GL(r, \mathbb{R}) \quad \text{or} \quad \rho : G \mapsto GL(r, \mathbb{C}),$$

from G to a group of real or complex matrices.

Definition (Character of a Representation): Given a representation ρ of G, the map

$$\chi_p : \; G \mapsto \mathbb{R} \quad \text{or} \quad G \mapsto \mathbb{C},$$

defined by

$$\chi_\rho(g) = \mathrm{tr}\rho(g), \qquad g \in G$$

is called the *character of the representation* ρ.

Definition (Irreducibility): A representation ρ is called irreducible if the vector space $\mathbb{R}^r$ or $\mathbb{C}^r$ contains no proper subspaces that are invariant under the matrix group ρ.

Definition (G-invariance): A subspace Q of the representation space $\mathbb{R}^r$ is said to be invariant under the matrix group $\rho(G)$ or *G-invariant* if

$$\rho(g)W \subset W, \qquad \forall g \in G.$$

Then we can restrict ρ to W and have a sub-representation of ρ.

Lemma (Schur's): Let

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

be two irreps (irreducible representations) of a group G. If $A : \mathbb{R}^{r_1} \to \mathbb{R}^{r_2}$ is a linear transformation changing ρ_1 into ρ_2 as

$$A\rho_1(g) = \rho_2(g)A, \qquad \forall g \in G$$

then A is either trivial or a bijection in which case $r_1 = r_2$.

If G is a Lie group and a representation taking G into the space of matrices ρ is smooth, then the push-forward map ρ_* is a linear map from the Lie algebra $\mathcal{G} = T^{(1)}$ to the space of all $r \times r$ matrices

$$\rho_* : \mathcal{G} \to Mat(r, \mathbb{R}). \tag{3.9}$$

Moreover ρ_* is a representation of the Lie algebra since it is a homomorphism which is linear and preserves the commutator operations

$$\rho_* \underbrace{[\xi, \eta]}_{\text{Lie Bracket}} = \underbrace{[\rho_* \xi, \rho_* \eta]}_{\text{Matrix Commutator}} .$$

Definition (Faithful Representation): A representation $\rho : G \to GL(r, \mathbb{R} \text{ or } \mathbb{C})$ is called *faithful* if it is one-to-one. In other words, if its kernel is trivial, $\rho(g) \neq \mathbb{I}$ unless $g = g_0$. Any matrix Lie group has a faithful representation and any Lie group with a faithful representation can be realised as a matrix Lie group.

An example of when this would not be possible would be for the group $\tilde{SL}(2, \mathbb{R})$ of all transformations of the real line of the form

$$x \mapsto x + 2\pi a + -i \ln \left(\frac{1 - e^{-ix}}{1 - \bar{z} e^{ix}} \right)$$

where $a, x \in \mathbb{R}$, $z \in \mathbb{C}$ with $|z| < 1$ and the logarithm is chosen as the main branch.

3.4 Adjoint Representation

Definition (Inner Automorphism): For each $h \in G$ the transformation defined by $g \mapsto hgh^{-1}$ is called the *inner automorphism* of G determined by h. The identity is invariant under this map and so the push-forward map of the tangent space T at $g_0 \in G$ is a linear transformation of T denoted by $\mathrm{Ad}_h : T \to T$.

It satisfies

1. $\mathrm{Ad}_{g_0} = \mathbb{I}$.
2. $\mathrm{Ad}_{h_1} \mathrm{Ad}_{h_2} = \mathrm{Ad}_{h_1 h_2}$ for all $h_1, h_2 \in G$.
3. For a group element h with inverse h^{-1} we get $\mathrm{Ad}_{h^{-1}} = (\mathrm{Ad}_h)^{-1}$.

This means that the map $h \mapsto \mathrm{Ad}_h$ is a linear representation of the group G which we call the adjoint representation of G and it is defined by

$$\mathrm{Ad} : G \mapsto GL(n, \mathbb{R}), \qquad h \mapsto \mathrm{Ad}_h, \qquad n = \dim(G).$$

Let's take a quick look at what this looks like in local coordinates in a neighbourhood U_0. We shall denote the inner automorphism of G determined by h as $\mathrm{AD}(h) : g \to hgh^{-1}$. The corresponding push-forward map is

$$AD(h)_* : \xi^i \to \eta^i = \frac{\partial \psi^i(\psi(h,x),\varphi(h))}{\partial x^j}\xi^j, \qquad \xi \in T_g G, \qquad \eta \in T_{hgh^{-1}}G$$

where we have specified $x = (x^1, \dots, x^n)$ as the local coordinates of g and left h alone for convenience.

There are a few facts that follow from these definitions in relation to one-parameter subgroups. Let $F(t) = \exp At$ be a one-parameter subgroup of a Lie group G. Then:

1. $\mathrm{Ad}_{F(t)}$ is a one-parameter subgroup of $GL(n,\mathbb{R})$.

2. The vector $\partial_t \mathrm{Ad}_{F(t)}|_{t=0}$ lies in the Lie algebra $\mathfrak{gl}(n,\mathbb{R}) \cong \mathrm{Mat}(n,\mathbb{R})$ of the group $GL(n,\mathbb{R})$ and can be regarded as a linear operator. This operator is denoted ad_A and is given by

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

This formula is obtained using the push forward map defined above with the replacements $g = g_0 \iff x = 0$, $\xi, \eta \in T$ and working in the local coordinates of the group element $h \to F(t) = (f^1(t), \dots, f^n(t))$ where we can take advantage of the fact that $\dot{f}^i(0) = A$.

Definition (Simple Lie Algebras): A Lie algebra $\mathcal{G}$ is said to be *simple* if

1. It is *not* commutative.

2. It has *no proper ideals*, which are non-trivial subspaces $\mathcal{I} \neq \mathcal{G}$ for which $[\mathcal{I},\mathcal{G}] \subset \mathcal{I}$.

Or is said to be *semi-simple* if

$$\mathcal{G} = \mathcal{I}_1 \oplus \cdots \oplus \mathcal{I}_k$$

where $\mathcal{I}_j$ are ideals which are simple as Lie algebras. These ideals are mutually commutative for $i \neq j$. A Lie group is semi-simple or simple if its algebra is.

Definition (Killing Form): The *Killing form* on an arbitrary Lie algebra $\mathcal{G}$ is the binary operation defined by

$$\langle A, B \rangle = -\mathrm{tr}(ad_A ad_B). \tag{3.10}$$

Theorem (Adjoint Rep of a Simple Group is an Irrep): If the Lie algebra $\mathcal{G}$ of a group G is simple, then the linear representation Ad: $G \to GL(n,\mathbb{R})$ is *irreducible.* And in this context irreducible means it has no proper invariant subspaces under the group of inner automorphisms Ad_G.

Theorem (Semi-Simple Lie Algebras have Positive Definite Killing Forms): If the Killing form of a Lie algebra is positive definite the the Lie algebra is semi-simple.

There is an even stronger theorem than this one which says that a Lie algebra is semi-simple if the Killing form is non-degenerate.

Definition (Representation by Left Action): We say that a Lie group G is represented as a group of transformations of a manifold M, or has a *left action* on M if:

1. There is a diffeomorphism from M to itself associated with each of its elements g

$$x \mapsto \mathcal{T}_g(x), \qquad x \in M$$

   such that $\mathcal{T}_{gh} = \mathcal{T}_g \mathcal{T}_h$ for all $g, h \in G$.

2. Each $\mathcal{T}_g(x)$ depends smoothly on the arguments g, x making $(g, x) \mapsto \mathcal{T}_g(x)$ a smooth map.

The Lie group is said to have a *right action* if the above definition is valid with the alternative product of diffeomorphisms $\mathcal{T}_g \mathcal{T}_h = \mathcal{T}_{hg}$.

An example of such a representation would be for any linear Lie group of real $n \times n$ matrices acting on $\mathbb{R}^n$. There are more manifolds that we will need to consider so to make sure we are ready to deal with any of them we need some more definitions.

Definition (Transitive Group Action): The action of G on M is said to be *transitive* if for every two points $x, y \in M$, there exists an element g of G such that $\mathcal{T}_g(x) = y$. In other words, if we can go between any two points on the manifold by the action of some group element.

In practice it is sufficient to choose a reference point $x_0 \in M$ as the origin and prove that for a general y there exists a group element g such that $\mathcal{T}_g(x_0) = y$.

Definition (Homogeneous Space): A manifold on which a Lie group acts transitively is called a *homogeneous space* of the Lie group. The Lie group itself trivially satisfies this definition since $T_g(h) = gh \in G$. The manifold of the Lie group is called the *principal* homogeneous space.

Definition (Isotropy Group): Let x be any point of a homogeneous space M of a Lie group G. The *isotropy* group H_x of the point x is the stabiliser of x under the action of G, which is the subgroup of elements of G under the action of which x is invariant.

$$H_x = \{h \mid \mathcal{T}_h(x) = x\} \tag{3.11}$$

Lemma (Isotropy Groups are Isomorphic): All isotropy groups H_x of points x of a homogeneous space are isomorphic to each other.

Theorem (Correspondence between Hom. Spaces and Left Cosets): There is a one-to-one correspondence between the points of a homogeneous space M of a group G and the left cosets gH of H in G where H is the isotropy group and G acts on the left.

4 Vector Bundles on a Manifold

4.1 The Tangent Bundle

Definition (Tangent Bundle): The *tangent bundle* $T(M)$ of an n-dimensional manifold is a $2n$-dimensional manifold defined as follows:

1. The points of $T(M)$ are the pairs (x, ξ) where $x \in M$ and $\xi \in T_xM$.
2. Given a chart U_q of M with the local coordinates (x_q^i), the corresponding charts U_q^T of $T(M)$ is the set of all pairs (x, ξ) where

$$x = (x_q^1, \ldots, x_q^n) \in U_q \qquad \text{and} \qquad \xi = \xi_q^i \frac{\partial}{\partial x_q^i} \in T_xM,$$

with local coordinates

$$(y_q^1, \ldots, y_q^{2n}) = (x_q^1, \ldots, x_q^n, \xi_q^n, \ldots, \xi_q^1) = (x_q^i, \xi_q^i).$$

Proposition (T(M) is Smooth and Oriented): The tangent bundle $T(M)$ is a smooth, oriented $2n$-dimensional manifold.

The above proposition is easily proven by showing that the Jacobian matrix of the transition functions is non-degenerate.

4.2 The Cotangent Bundle

Definition (Cotangent Bundle): The *cotangent bundle* $T^*(M)$ of an n-dimensional manifold M is a $2n$-dimensional manifold defined as follows:

1. The points of $T^*(M)$ are the pairs (x, p), where $x \in M$ and p is a co-vector at the point x, so $p \in T_x^*M$.
2. Given a chart U_q of M with local coordinates (x_q^i), the corresponding chart $U_q^{T^*}$ of $T^*(M)$ is the set of all pairs (x, p) where

$$x = (x_q^1, \ldots, x_q^n) \in U_q \qquad \text{and} \qquad p = (p_q)_i dx^i \in T_x^*M$$

with local coordinates

$$y = (y_q^1, \ldots, y_q^{2n}) = (x_q^1, \ldots, x_q^n, p_{q1}, \ldots, p_{qn}) = (x_q^i, p_{qi}).$$

So we can make a clear correlation to the case we saw in classical mechanics where the tangent bundle is the set of points which are coordinates and their velocity vectors while a cotangent bundle is the set of points with their conjugate momenta.

Proposition ($T^*(M)$ is Smooth and Oriented): The cotangent bundle is a smooth, oriented $2n$-dimensional manifold.

As a final note on this short section, we note that the existence of a metric g_{ij} on the manifold M gives rise to a map from the tangent bundle to the cotangent bundle

$$T(M) \mapsto T^*(M): \quad (x^i, \xi^i) \mapsto (x^i, g_{ij}\xi^j) \tag{4.1}$$

If we define the differential one form $\omega = p_i dx^i$ on M, we know it is invariant under a change of coordinates of $T^*(M)$ so it is a differential form on $T^*(M)$. Its differential

$$d\omega = dp_i \wedge dx^i$$

is *a non-degenerate, closed,* $(d(d\omega) = 0)$ 2-form on $T^*(M)$ and therefore $T^*(M)$ is a symplectic manifold.

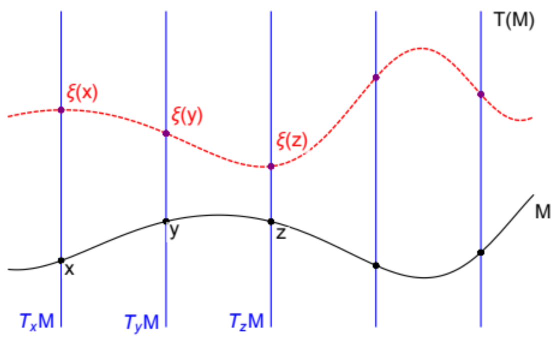
5 Vector Fields and Lie Derivatives

5.1 Vector and Tensor Fields

Definition (Vector Field): A *vector field* is a map that specifies a unique vector at each point x of a manifold M

$$\xi : M \to T(M), \qquad x \mapsto \xi_x \in T_x M \tag{5.1}$$

In other words a vector field intersects each tangent space in the whole bundle $T(M)$ at *only one point* by taking a cross section of $T(M)$. A nice picture which describes this definition is shown below:



In a specific coordinate system on the manifold (x_p^i) we can write

$$\xi = \xi_p^i(x)\frac{\partial}{\partial x_p^i}, \qquad x \in U_p \tag{5.2}$$

and the subscript p can actually be dropped since there exists a unique vector at each point. In this form, a vector field can be understood as a differential operator which maps a scalar function to a scalar function *on* M via the equation

$$\xi(f) = \xi^i\frac{\partial f}{\partial x^i}, \quad \Longrightarrow \quad \xi(f) : M \to \mathbb{R}, \qquad x \mapsto \xi^i(x)\frac{\partial f(x)}{\partial x^i}. \tag{5.3}$$

Lastly, vector fields are linear maps and they satisfy the Leibniz rule

$$\xi(f,g) = \xi(f)g + f\xi(g). \tag{5.4}$$

Now we can generalise this definition to objects of more general rank, i.e. tensors.

Definition (Tensor Field): A *tensor field* of type (r,s) assigns a unique tensor of type (r,s) to each point x of the manifold M

$$^{(r,s)}\xi : M \mapsto T^{(r,s)}(M), \qquad x \mapsto^{(r,s)} \xi_x \in T_x^{(r,s)}M$$

so similar to a vector field, it is a cross section of $T^{(r,s)}(M)$. A co-vector field or a field of one forms is a cross section of the cotangent bundle $T^*(M)$.

Aside (Motivation for the Lie Bracket): Consider the following composition of vector fields

$$\xi(\eta(f)) = \xi^i\frac{\partial}{\partial x^i}\left(\eta^j\frac{\partial f}{\partial x^j}\right) = \xi^i\frac{\partial \eta^j}{\partial x^i}\frac{\partial f}{\partial x^j} + \xi^i\eta^j\frac{\partial^2 f}{\partial x^i \partial x^j} \tag{5.5}$$

which is not a vector field since the second derivative wouldn't satisfy the Leibniz rule.

Definition (Lie Bracket): The commutator or *Lie bracket* defined by

$$[\xi,\eta](f) \equiv \xi(\eta(f)) - \eta(\xi(f)) = \left(\xi^i\frac{\partial \eta^j}{\partial x^i} - \eta^i\frac{\partial \xi^j}{\partial x^i}\right)\frac{\partial f}{\partial x^j}, \tag{5.6}$$

is a vector field with components given by the terms inside the bracket above. The bracket satisfies

1. **Anti-Commutative**: $[\xi,\eta] = -[\eta,\xi]$.
2. **Linear:** $[\xi,\eta+\zeta] = [\xi,\eta] + [\xi,\zeta]$.
3. **Leibniz Rule:** $[\xi, f\eta] = f[\xi,\eta] + \xi(f)\eta$.
4. **Jacobi Identity:** $[\xi,[\eta,\zeta]] + [\eta,[\zeta,\xi]] + [\zeta[\xi,\eta]] = 0$.

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

5.2 Integral Curves

Definition (Integral Curves): Let $\xi^i(x)$ be a vector field on M. Consider this autonomous system of differential equations

$$\dot{x}(t) \equiv \frac{dx^i}{dt} = \xi^i(x^1(t), \ldots, x^n(t)) \tag{5.7}$$

which are called autonomous because they have no explicit dependence on t. The solutions $x^i = x^i(t)$ to this system are called the *integral curves* of the vector field ξ^i. Then the vector field $\xi^i(x)$ is comprised of tangent vectors to the integral curves. We denote by

$$F_t^i(x_0^1, \ldots, x_0^n) = x^i = x^i(t, x_0^1, \ldots, x_0^n) \tag{5.8}$$

the integral curve of the vector field ξ^i satisfying the initial condition

$$x^i|_{t=t_0} = x^i \tag{5.9}$$

and with the conventional choice $t_0 = 0$ we use the notation

$$x_0 = x_0^1, \ldots, x_0^n. \tag{5.10}$$

The formula $F_t^i(x_0^1, \ldots, x_0^n) = x^i$ defines a self-map

$$F_t : x_0 = (x_0^1, \ldots, x_0^n) \mapsto (x^1(t, x_0), \ldots, x^n(t, x_0)) \tag{5.11}$$

of our manifold, dependent on the parameter t. In mechanics terms, F_t applied to a point $x_0 \in M$ gives the position of the particle after a time t as the particle moves along the integral curve through x_0.

Definition (Local Group): From the theory of differential equations, given any point x_0 at which $(\xi^i) \neq 0$ the map F_t is *locally* a diffeomorphism , i.e. there exists a neighbourhood of x_0 on which, for sufficiently small t, F_t is a diffeomorphism. Then for sufficiently small values of the parameters t and s, the diffeomorphisms satisfy

$$F_{t+s} = F_t \circ F_s = F_s \circ F_t, \qquad F_{-t} = (F_t)^{-1}. \tag{5.12}$$

Then we say that the diffeomorphsims F_t define a *local group.*

Definition (Flow of a Vector Field): This local abelian one-parameter group of diffeomorphisms F_t defined above is called the *flow* generated by the vector field ξ^i.

For small values t we can find a solution to the defining equations of the integral curves (5.7) since we can expand the equation of flow in powers of t as a Taylor series:

$$x^i(t, x_0) = x_0^i + t\frac{dx^i}{dt}(0) + \frac{1}{2}t^2\frac{d^2x^i}{dt^2} + \mathcal{O}(t^3). \tag{5.13}$$

Then using that

$$\frac{d^2x^i}{dt^2} = \frac{d}{dt}\xi^i(x^1(t),\dots,x^n(t)) = \frac{\partial\xi^i}{\partial x^j}\frac{dx^j}{dt} = \frac{\partial\xi^i}{\partial x^j}\xi^j(x(t)).$$

We get

$$x^i(t,x_0) = x_0^i + t\xi^i(x_0) + \frac{1}{2}t^2\frac{\partial\xi^i}{\partial x^j}\xi^j(x_0) + \mathcal{O}(t^3) \tag{5.14}$$

so that the Jacobian matrix and its inverse are

$$\frac{\partial x^i}{\partial x_0^j} = \delta_j^i + t\frac{\partial\xi^i}{\partial x_0^j} + \mathcal{O}(t^2) \qquad \text{with inverse} \qquad \frac{\partial x_0^i}{\partial x^j} = \delta_j^i - t\frac{\partial\xi^i}{\partial x^j} + \mathcal{O}(t^2). \tag{5.15}$$

This construction can of course be reversed. Given a one-parameter local group of diffeomorphisms $F_t = (F_t^1,\dots,F_t^n)$ we define its *velocity field* to be the vector field

$$\xi^i = \left(\frac{d}{dt}F_t^i\right)\bigg|_{t=0}, \qquad i = 1,\dots,n. \tag{5.16}$$

There is a way to interpret the Lie bracket or commutator of vector fields geometrically through the idea of flow. Let ξ and η be vector fields on M with which we associate the flows F_t and G_t respectively. In general the two flows are non-commuting $G_s\circ F_t \neq F_t\circ G_s$ but to get to a point $(G_s\circ F_t)(x)$ from the point x on the manifold we move to $F_t(x)$ along the integral curve of ξ and then from there to $G_s(F_t(x))$ along the integral curve of η. Now using (5.14) we can derive the commutator of these flows by computing the Taylor expansion for small t,s to quadratic order.

$$(G_s\circ F_t(x))^i = G_s^i(F_t(x)) = F_t^i(x) + s\eta^i(F_t(x)) + \frac{1}{2}s^2\frac{\partial\eta^i}{\partial x^j}\eta^j(F_t(x))$$

$$x^i + t\xi^i + \frac{1}{2}t^2\frac{\partial\xi^i}{\partial x^j}\xi^j(x) + s\eta^i(x+t\xi(x)) + \frac{1}{2}s^2\frac{\partial\eta^i}{\partial x^j}\eta^j(x)$$

$$x^i + t\xi^i(x) + \frac{1}{2}t^2\frac{\partial\xi^i}{\partial x^j}\xi^j(x) + \frac{1}{2}s^2\frac{\partial\eta^i}{\partial x^j}\eta^j(x) + s\eta^i(x^i) + st\frac{\partial\eta^i}{\partial x^j}\xi^j(x).$$

Commuting operations simply gives the same thing with the vector fields playing opposite roles

$$(F_t\circ G_s(x))^i = (G_s\circ F_t(x))^i|_{\xi\longleftrightarrow\eta,\qquad t\longleftrightarrow s}$$

which gives us that the commutator of the two flows for small t,s is

$$[G_s,F_t](x) = G_s(F_t(x)) - F_t(G_s(x)) = ts[\xi,\eta] + \mathcal{O}(t^3,s^3). \tag{5.17}$$

This intuitively measures the discrepancy between the points obtained by following the integral curves of ξ and η in different orders. The vectors comprising a coordinate-induced basis must commute because the partial derivatives do and conversely if all elements of a basis for vector fields commute then the basis is coordinate-induced.

Definition (Action of $F_t(x)$ on Smooth Functions): A local group $F_t(x)$ of diffeomorphisms $F_t(x)$ with an associated vector field $\xi(x)$ is defined the act on smooth functions f as follows

$$(F_t f)(x) = f(F_t(x)). \tag{5.18}$$

The simplest example would be a group of translations $F_t(x) = x + t$ in $\mathbb{R}$ such that $(F_t f)(x) = f(x + t)$.

Definition (Exponential Function of a Vector Field): The exponential function of a vector field ξ is the operator

$$e^{t\partial_\xi} = 1 + t\partial_\xi + \frac{t^2}{2}(\partial_\xi)^2 + \cdots = \sum_{k=0}^{\infty} \frac{t^k}{k!}(\partial_\xi)^k \tag{5.19}$$

where $\partial_\xi \equiv \xi^i \frac{\partial}{\partial x^i}$ is the directional derivative operator along ξ. The action of $e^{t\partial_\xi}$ on functions $f(x)$ is defined by

$$e^{t\partial_\xi} f(x) = f(x) + t\partial_\xi f(x) + \frac{t^2}{2}(\partial_\xi)^2 f(x) + \cdots. \tag{5.20}$$

Proposition (Equivalence of $F_t(x)$ Action and $e^{t\partial_\xi}$ Action): For analytic vector fields $\xi(x)$ and analytic functions $f(x)$, the exponential function of $\xi(x)$ coincides with the action of the local group diffeomorphisms $F_t(x)$ on the smooth functions f for sufficiently small t.

5.3 The Lie Derivative

Now we want to define an operator based on the flow F_t generated by a vector field ξ which acts on tensors of any general type. Naively defining the action on tensors as an operation which evaluates the tensors at different points along the flow would be incorrect since we do not know how to compare tensors at different points of a manifold yet. Instead, we picture it as a passive transformation which acts on the coordinate basis vectors by $F_t^{-1} = F_{-t}$.

Definition (Action of $F_t(x)$ on Tensors): A one-parameter group of diffeomorphisms $F_t(x)$ with an associated vector field $\xi(x)$ is defined to *act on smooth tensors* $T = (T^{i_1,\dots,i_p}_{j_1,\dots,j_q})$ of type (p, q) as

$$(F_t T)^{i_1 \dots i_p}_{j_1 \dots j_q}(x) = T^{k_1 \dots k_p}_{l_1 \dots l_q}(y) \frac{\partial y^{l_1}}{\partial x^{j_1}} \cdots \frac{\partial y^{l_q}}{\partial x^{j_q}} \frac{\partial x^{i_1}}{\partial y^{k_1}} \cdots \frac{\partial x^{i_p}}{\partial y^{k_p}} \tag{5.21}$$

where $y^i = F_t^i(x)$. It is clear in this definition that $(F_t T)^{i_1 \dots i_p}_{j_1,\dots,j_q}$ and $T^{k_1 \dots k_p}_{l_1 \dots l_q}$ are the components of the same tensor measured in different coordinate systems which differ by the transformation of the basis vectors along the flow.

Having an action on tensors as above we can define a type of derivative which is taken on tensors in the direction of a vector field.

Definition (Lie Derivative of a Tensor): The Lie derivative of a tensor $T = (T^{i_1 \dots i_p}_{j_1 \dots j_q})$ along a vector field ξ is the tensor $L_\xi T$ given by

$$L_\xi T^{i_1 \dots i_p}_{j_1 \dots j_q} = \left[\frac{d}{dt}(F_t T)^{i_1 \dots i_p}_{j_1 \dots j_q}\right]. \tag{5.22}$$

Regarding the diffeomorphism F_t as a time-dependent deformation of the manifold M, the Lie derivative is said to measure the rate of change of the tensor T resulting from this deformation. To find an explicit formula for the Lie derivative here we expand $F_t T$ up to linear order in t. First we use the conventional coordinate definitions

$$y^i = F^i_t(x) = x^i + t\xi^i(x) + \mathcal{O}(t^2), \qquad \frac{\partial y^i}{\partial x^j} = \delta^i{}_j + t\frac{\partial \xi^i}{\partial x^j}(x) + \mathcal{O}(t^2).$$

$$\frac{\partial x^i}{\partial y^j} = \delta^i{}_j - t\frac{\partial \xi^i}{\partial x^j}(x) + \mathcal{O}(t^2).$$

For the expansion of components $(F_t T)^{i_1,\dots,i_p}_{j_1,\dots,j_q}$ we then obtain

$$\begin{aligned}
(F_t T)^{i_1 \dots i_p}_{j_1 \dots j_q}(x) &= T^{k_1 \dots k_p}_{l_1 \dots l_q}(x + t\xi)\left(\delta^{l_1}_{j_1} + t\frac{\partial \xi^{l_1}}{\partial x^{j_1}}\right)\cdots\left(\delta^{l_q}_{j_q} + t\frac{\partial \xi^{l_q}}{\partial x^{j_q}}\right) \\
&\times\left(\delta^{i_1}_{k_1} - t\frac{\partial \xi^{i_1}}{\partial x^{k_1}}\right)\cdots\left(\delta^{i_p}_{k_p} - t\frac{\partial \xi^{i_p}}{\partial x^{k_p}}\right) \\
&= T^{i_1 \dots i_p}_{j_1 \dots j_q}(x) + t\Bigg[\xi^a \frac{\partial T^{i_1 \dots i_p}_{j_1 \dots j_q}}{\partial x^a}(x) \\
&+T^{i_1 \dots i_p}_{a j_2 \dots j_q}\frac{\partial \xi^a}{\partial x^{j_1}} + T^{i_1 \dots i_p}_{j_1 a j_3 \dots j_q}\frac{\partial \xi^a}{\partial x^{j_2}} + \cdots + T^{i_1 \dots i_p}_{j_1 \dots j_{q-1} a}\frac{\partial \xi^a}{\partial x^{j_q}} \\
&-T^{a i_2, \dots i_p}_{j_1 \dots j_q}\frac{\partial \xi^{i_1}}{\partial x^a} - T^{i_1 a i_3 \dots i_p}_{j_1 \dots j_q}\frac{\partial \xi^{i_2}}{\partial x^a} - \cdots - T^{i_1 \dots i_{p-1} a}_{j_1 \dots j_q}\frac{\partial \xi^{i_p}}{\partial x^a}\Bigg]
\end{aligned}$$

which gives the explicit form for $L_\xi T^{i_1 \dots i_p}_{j_1 \dots j_q}$ as the term in square brackets.

Theorem (Lie Bracket of Vector Fields Aligned with Coordinates on $\mathbb{R}^n$): Let $\xi_1, \dots, \xi_m$ be vector fields in $\mathbb{R}^n$. If there is a system of coordinates $y^1, \dots, y^n$ such that at every point the vector ξ_j is tangent to the axis of y^j then the fields satisfy

$$[\xi_j, \xi_k] = f^{(1)}_{jk}\xi_j + f^{(2)}_{jk}\xi_k \tag{5.23}$$

where $f^{(a)}_{jk}$ are scalars.

The Lie derivative definition is quite involved and index heavy so it is useful to look at its form for tensors of type $(0,1)$ and $(0,2)$. If $T = T_i$ is a co-vector then

$$L_\xi T_i = \xi^a \frac{\partial T_i}{\partial x^a} + T_a \frac{\partial \xi^a}{\partial x^i}. \tag{5.24}$$

In differential notation, if T_i can be written as a 1-form $\frac{\partial f}{\partial x^i} \equiv df_i$, then

$$L_\xi T_i = L_\xi df_i = \xi^a \frac{\partial^2 f}{\partial x^a \partial x^i} + \frac{\partial f}{\partial x^a}\frac{\partial \xi^a}{\partial x^i} = \frac{\partial}{\partial x^i}\left(\xi^a \frac{\partial f}{\partial x^a}\right) = \frac{\partial}{\partial x^i}(L_\xi f). \tag{5.25}$$

This reveals a deeper property of the Lie derivative in that it commutes with the differential operator d:

$$L_\xi df = d(L_\xi f). \tag{5.26}$$

For the tensor of type $(0,2)$ we will immediately specialise to a metric $T = g_{ij}$. Which has the Lie derivative

$$L_\xi g_{ij} = \xi^a \frac{\partial g_{ij}}{\partial x^a} + g_{aj}\frac{\partial \xi^a}{\partial x^j} + g_{ia}\frac{\partial \xi^a}{\partial x^j} \equiv u_{ij}. \tag{5.27}$$

The tensor u_{ij} is called the *strain tensor* and it describes how the metric g_{ij} changes for small deformations F_t defined by the vector field ξ.

Definition (Killing Vector): If $L_\xi g_{ij} = 0$ for some metric g_{ij} of a space, then ξ is called a *Killing vector*. This means for example that if (y^i) is a coordinate system and $\xi = (1, 0, \ldots, 0)$ a Killing vector field, then the metric g_{ij} is independent of y^1.

Lemma (Killing Vector Fields of a Riemannian Manifold): The Killing vector fields of a (pseudo-) Riemannian manifold form a Lie algebra with respect to the Lie bracket given by the commutator.

Proof: The proof is nice and simple, if ξ and η are two Killing vectors, then

$$L_{[\xi,\eta]} g_{ij} = [L_\xi, L_\eta] g_{ij} = 0 \tag{5.28}$$

which means that $[\xi, \eta]$ is also a Killing vector.

5.4 The Lie Derivative of a Volume Element

Now we can see how the Lie derivative behaves when it operates on a volume element which is a tensor defined by $T = \sqrt{|g|}\epsilon_{i_1 \ldots i_n}$ where g denotes the determinant of the metric tensor as usual. The Lie derivative yields

$$L_\xi(\sqrt{|g|}\epsilon_{i_1\ldots i_n}) = \xi^a \frac{\partial \sqrt{|g|}}{\partial x^a}\epsilon_{i_1 \ldots i_n} + \sqrt{|g|}\left(\epsilon_{a i_2 \ldots i_n}\frac{\partial \xi^a}{\partial x^{i_1}} + \cdots + \epsilon_{i_1 \ldots i_{n-1} a}\frac{\partial \xi^a}{\partial x^{i_n}}\right) \tag{5.29}$$

We can define the Levi-Civita symbol to be either 0 or 1 for the multi-index $(i_1 \dots i_n)$ and that then determines the form of the Lie derivative. In each case:

$$(1) \qquad \epsilon_{i_1 \dots i_n} \neq 0 \quad \Longrightarrow \quad \epsilon_{a i_2 \dots i_n} \frac{\partial \xi^a}{\partial x^{i_1}} + \dots + \epsilon_{i_1 \dots i_{n-1} a} \frac{\partial \xi^a}{\partial x^{i_n}} \tag{5.30}$$

$$= \epsilon_{i_1 \dots i_n} \left(\frac{\partial \xi^{i_1}}{\partial x^{i_1}} + \dots + \frac{\partial \xi^{i_n}}{\partial x^{i_n}} \right) = \epsilon_{i_1 \dots i_n} \frac{\partial \xi^a}{\partial x^a}.$$

$$(2) \qquad \epsilon_{i_1 \dots i_n} = 0 \quad \Longrightarrow \quad \epsilon_{a i_2 \dots i_n} \frac{\partial \xi^a}{\partial x^{i_1}} + \dots + \epsilon_{i_1 \dots i_{n-1} a} \frac{\partial \xi^a}{\partial x^{i_n}} = 0. \tag{5.31}$$

Thus we use the first case since it is non-trivial and then the Lie derivative is

$$L_\xi (\sqrt{|g|} \epsilon_{i_1 \dots i_n}) = \xi^a \frac{\partial \sqrt{|g|}}{\partial x^a} \epsilon_{i_1 \dots i_n} + \sqrt{|g|} \epsilon_{i_1 \dots i_n} \frac{\partial \xi^a}{\partial x^a}.$$

Or more compactly

$$L_\xi \sqrt{|g|} \epsilon_{i_1 \dots i_n} = \epsilon_{i_1 \dots i_n} \frac{\partial}{\partial x^a} (\sqrt{|g|} \xi^a). \tag{5.32}$$

This means that in the special case where g_{ij} is the metric on an oriented, closed manifold, the volume does not change under small deformations F_t defined by a vector field ξ. We identify a normalised version of the derivative in (5.32) as the divergence of the vector field ξ defined as

$$\nabla_a \xi^a \equiv \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^a} (\sqrt{|g|} \xi^a). \tag{5.33}$$

Lastly, using the identity

$$\frac{\partial g}{\partial x^a} = g g^{ij} \frac{\partial g_{ij}}{\partial x^a} \iff \frac{\partial \ln |g|}{\partial x^a} = g^{ij} \frac{\partial g_{ij}}{\partial x^a}$$

we can find a relation to the strain tensor u_{ij} as follows. We have on one hand

$$L_\xi (\sqrt{|g|} \epsilon_{i_1 \dots i_n}) = \xi^a \frac{\partial \sqrt{|g|}}{\partial x^a} \epsilon_{i_1 \dots i_n} + \sqrt{|g|} \epsilon_{i_1 \dots i_n} \frac{\partial \xi^a}{\partial x^a}$$

$$= \sqrt{|g|} \epsilon_{i_1 \dots i_n} \left(\xi^a \frac{1}{2} g^{ij} \frac{\partial g_{ij}}{\partial x^a} + \frac{\partial \xi^a}{\partial x^a} \right).$$

While on the other hand

$$g^{ij} u_{ij} = g^{ij} L_\xi g_{ij} = g^{ij} \left(\xi^a \frac{\partial g_{ij}}{\partial x^a} + g_{aj} \frac{\partial \xi^a}{\partial x^j} + g_{ia} \frac{\partial \xi^a}{\partial x^j} \right) = \xi^a g^{ij} \frac{\partial g_{ij}}{\partial x^a} + 2 \frac{\partial \xi^a}{\partial x^a}.$$

Therefore we have obtained the relations

$$L_\xi \sqrt{|g|} \epsilon_{i_1 \dots i_n} = \frac{1}{2} g^{ij} u_{ij} \sqrt{|g|} \epsilon_{i_1 \dots i_n}, \tag{5.34}$$

$$\nabla_a \xi^a = \frac{1}{2} g^{ij} u_{ij}. \tag{5.35}$$

6 Covariant Differentiation

Our aim in the following analysis is to find some differentiation operations on tensors which again gives us tensors instead of scalars. The first example one might think of comes from our study of manifolds and that is the differential map.

6.1 The Differential Map on Tensors

The differential operator d transforms a skew-symmetric rank k tensor $T = (T_{i_1 \dots i_k})$ to another skew-symmetric tensor dT of rank $k+1$, explicitly

$$(dT)_{i_1 \dots i_k} = \sum_{q=1}^{k+1} (-1)^{q-1} \frac{\partial T_{i_1 \dots \hat{i}_q \dots i_{k+1}}}{\partial x^{i_q}} = \sum_{q=1}^{k+1} (-1)^{q-1} \partial_{i_q} T_{i_1 \dots \hat{i}_q \dots i_{k+1}}. \tag{6.1}$$

Where the index i_q is omitted by the hat. For $k = 1$, we have

$$(dT)_{ij} = \frac{\partial T_j}{\partial x^i} - \frac{\partial T_i}{\partial x^j} = \partial_i T_j - \partial_j T_i, \tag{6.2}$$

the combination is important here because $\partial_j T_i$ alone is *not* a tensor. The transformation of the partial derivative of the tensor component $\partial_k T_i$ is determined by the transformation between coordinates $x \to z$ given by

$$x^i = x^i(z^1, \dots, z^n), \qquad i = 1, \dots, n.$$

Then we obtain

$$\partial_q \tilde{T}_j = \frac{\partial \tilde{T}_j}{\partial z^q} = \frac{\partial}{\partial z^q}\left(T_i \frac{\partial x^i}{\partial z^j}\right) = \frac{\partial T_i}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i \frac{\partial^2 x^i}{\partial z^q \partial z^j} \tag{6.3}$$

$$\frac{\partial T_i}{\partial x^p}\frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i \frac{\partial^2 x^i}{\partial z^q \partial z^j}.$$

Thus

$$\partial_q \tilde{T}_j = \partial_p T_i \frac{\partial x^p}{\partial z^q}\frac{\partial x^i}{\partial z^j} + T_i \frac{\partial^2 x^i}{\partial z^q \partial z^j} \tag{6.4}$$

which is also not a tensor due to the second derivative term. The corresponding upper index rank-1 tensors transform as

$$\partial_q \tilde{T}^j = \frac{\partial \tilde{T}^j}{\partial z^q} = \frac{\partial}{\partial z^q}\left(T^i \frac{\partial z^j}{\partial x^i}\right) = \frac{\partial T^i}{\partial z^q}\frac{\partial z^j}{\partial x^i} + T^i \frac{\partial}{\partial z^q}\frac{\partial z^j}{\partial x^i}, \tag{6.5}$$

$$= \partial_p T^i \frac{\partial x^p}{\partial z^q}\frac{\partial z^j}{\partial x^i} + T^i \frac{\partial^2 z^j}{\partial x^p \partial x^i}\frac{\partial x^p}{\partial z^q}.$$

The Euclidean divergence $\partial_i T^i$ transforms as

$$\partial_j \tilde{T}^j = \partial_i T^i + T^i \frac{\partial^2 z^j}{\partial x^p \partial x^i}\frac{\partial x^p}{\partial z^j} \tag{6.6}$$

and it is *not a scalar.* We have not found exactly what we are looking for using the differential, but we are getting somewhere.

6.2 New Tensor Unlocked

Now lets return to tensors of type (q,p). We need quantities like $\partial_k T^{i_1 \dots i_p}_{j_1 \dots j_q}$ which transform as tensors (and therefore are tensors). In $\mathbb{R}^n$ we then introduce the tensor denoted by $\nabla_k T^{i_1 \dots i_p}_{j_1 \dots j_q}$ which is equal to $\partial_k T^{i_1 \dots i_p}_{j_1 \dots j_q}$ in Euclidean coordinates. We use some nice notation to write this tensor as

$$T^{i_1 \dots i_p}_{j_1 \dots j_q;\ k} \equiv \nabla_k T^{i_1 \dots i_p}_{j_1 \dots j_q} \equiv \nabla_k T^{(i)}_{(j)} = \partial_k T^{(i)}_{(j)} \tag{6.7}$$

To see how this tensor transforms, consider any other system with coordinates $z^1, \dots, z^n$. Then we obtain

$$\nabla_r \tilde{T}^{(k)}_{(l)} = \nabla_s T^{(i)}_{(j)} \frac{\partial x^s}{\partial z^r} \frac{\partial x^{(j)}}{\partial z^{(l)}} \frac{\partial z^{(k)}}{\partial x^{(i)}}. \tag{6.8}$$

where $(k) = k_1 \dots k_p$, $(l) = l_1 \dots l_q$, and so on. Now to see how we can expand equation (6.8) into its explicit form, consider the rank-1 tensors T_i, T^i which can be written in the z-coordinates as

$$\tilde{T}^k = T^i \frac{\partial z^k}{\partial x^i} \qquad \text{and} \qquad \tilde{T}_k = T_i \frac{\partial x^i}{\partial z^k},$$

with inverse transformations

$$T^i = \tilde{T}^k \frac{\partial x^i}{\partial z^k} \qquad \text{and} \qquad T_i = \tilde{T}_k \frac{\partial z^k}{\partial x^i}.$$

Using this we can expand the rank-1 version of (6.7) as

$$\nabla_r \tilde{T}^k = \partial_s T^i \frac{\partial x^s}{\partial z^r} \frac{\partial z^k}{\partial x^i} = \frac{\partial T^i}{\partial z^r} \frac{\partial z^k}{\partial x^i} = \frac{\partial}{\partial z^r} \left(T^i \frac{\partial z^k}{\partial x^i} \right) - T^i \frac{\partial}{\partial z^r} \frac{\partial z^k}{\partial x^i}$$

$$\implies \nabla_r \tilde{T}^k = \frac{\partial \tilde{T}^k}{\partial z^r} - \tilde{T}^s \frac{\partial x^i}{\partial z^s} \frac{\partial^2 z^k}{\partial x^m \partial x^i} \frac{\partial x^m}{\partial z^r}. \tag{6.9}$$

If we define the objects

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

then we can simply write

$$\nabla_r \tilde{T}^k = \frac{\partial \tilde{T}^k}{\partial z^r} + \Gamma^k_{sr} \tilde{T}^s. \tag{6.10}$$

Through the above analysis we have actually proven the following theorems.

Theorem (Transformation Formula of $\nabla_k T^i$): Let (T^i) be a vector field, and let $\nabla_k T^i = \frac{\partial T^i}{\partial x^k}$ be a tensor given in terms of Euclidean coordinates $x^1, \dots, x^n$. Then in arbitrary coordinates (z^r), $r = 1, \dots, n$, the transformed components of the tensor are given by:

$$\nabla_r \tilde{T}^k \equiv \tilde{T}^k{}_{;r} = \frac{\partial \tilde{T}^k}{\partial z^r} + \Gamma^k_{sr} \tilde{T}^s. \tag{6.11}$$

where the coefficients Γ^k_{sr} are defined as

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

Then we have an analogous result for the co-vector field T_i.

Theorem (Transformation Formula of $\nabla_k T_i$): Let (T_i) be a co-vector field, and let $\nabla_k T^i$ be a tensor given in terms of Euclidean coordinates $x^1, \ldots, x^n$ by $\nabla_k T_i = \frac{\partial T_i}{\partial x^k}$. Then in arbitrary coordinates (z^r), $r = 1, \ldots, n$, the transformed components of the tensor are given by:

$$\nabla_r \tilde{T}_k \equiv \tilde{T}_{k \;\; ;r} = \frac{\partial \tilde{T}_k}{\partial z^r} - \Gamma^k_{sr}\tilde{T}_s. \tag{6.12}$$

Now we can easily formulate such a theorem for general tensors by generalising the results above as follows.

Theorem (Transformation Formula of $\nabla_k T^{(i)}_{(j)}$): Let $T^{(i)}_{(j)}$ be the components of a tensor of type (p, q), and let $\nabla_k T^{(i)}_{(j)}$ be a tensor given in terms of the Euclidean coordinates $x^1, \ldots, x^n$ by the formula $\nabla_k T^{(i)}_{(j)} = \partial T^{(i)}_{(j)}/\partial x^k$. Then in arbitrary coordinates $z^1, \ldots, z^n$ the transformed components are given by the formula

$$\nabla_r \tilde{T}^{(k)}_{(l)} \equiv \tilde{T}^{(k)}_{(l);r} = \frac{\partial \tilde{T}^{(k)}_{(l)}}{\partial z^r} + \sum_{a=1}^{p} \Gamma^{k_a}_{sr} \tilde{T}^{k_1 \ldots (k_a \to s) \ldots k_p}_{l_1 \ldots l_q} - \sum_{a=1}^{q} \Gamma^{s}_{l_a r} \tilde{T}^{k_1 \ldots k_p}_{l_1 \ldots (l_a \to s) \ldots l_q}. \tag{6.13}$$

To make a general definition of this kind of tensor given by the ∇_r operation all that remains for us to figure out is how the coefficients Γ^k_{ij} transform under an arbitrary coordinate change. Consider going from $z \to z'$ coordinates. Then we can write down the tensors given by the co-vector $\tilde{T}_i$ in each system as

$$(z): \qquad \nabla_k \tilde{T}_i = \frac{\partial \tilde{T}_i}{\partial z^k} - \Gamma^r_{ik}\tilde{T}_r \qquad \text{and} \qquad (z'): \qquad \nabla_{k'} \tilde{T}_{i'} = \frac{\partial \tilde{T}_{i'}}{\partial z^{k'}} - \Gamma^{r'}_{i'k'}\tilde{T}_{r'}. \tag{6.14}$$

Since they are tensors we can transform the unprimed one and compare the expressions:

$$\nabla_{k'} \tilde{T}_{i'} = \nabla_k \tilde{T}_i \frac{\partial z^i}{\partial z^{i'}}\frac{\partial z^k}{\partial z^{k'}} = \left(\partial_k \tilde{T}_i - \Gamma^r_{ik}\tilde{T}_r\right)\frac{\partial z^i}{\partial z^{i'}}\frac{\partial z^k}{\partial z^{k'}} \tag{6.15}$$

On the other hand if we expand the expression for the primed tensor we get

$$\nabla_{k'} \tilde{T}_{i'} = \frac{\partial \tilde{T}_{i'}}{\partial z^{k'}} - \Gamma^{r'}_{i'k'}\tilde{T}_{r'} = \partial_k \tilde{T}_i \frac{\partial z^i}{\partial z^{i'}}\frac{\partial z^k}{\partial z^{k'}} + \tilde{T}_r \frac{\partial^2 z^r}{\partial z^{i'} \partial z^{k'}} - \Gamma^{r'}_{i'k'}\tilde{T}_r \frac{\partial z^r}{\partial z^{r'}} \tag{6.16}$$

where we have transformed $\partial_{k'} \tilde{T}_i$ using equation (6.4). Comparing these two expressions we get

$$\Gamma^{r'}_{i'k'}\frac{\partial z^r}{\partial z^{r'}} = \Gamma^r_{ik}\frac{\partial z^i}{\partial z^{i'}}\frac{\partial z^k}{\partial z^{k'}} + \frac{\partial^2 z^r}{\partial z^{i'} \partial z^{k'}}. \tag{6.17}$$

Multiplying by $\partial z^{s'}/\partial z^r$ and summing over r we get

$$\Gamma^{r'}_{i'k'} = \Gamma^r_{ik}\frac{\partial z^i}{\partial z^{i'}}\frac{\partial z^k}{\partial z^{k'}}\frac{\partial z^{s'}}{\partial z^r} + \frac{\partial^2 z^r}{\partial z^{i'}\partial z^{k'}}\frac{\partial z^{s'}}{\partial z^r}.$$

Finally we are ready to define what the operation actually is that produces these tensors.

Definition (Covariant Differentiation): An operation of covariant differentiation of tensors is said to be defined if we are given in terms of any system of coordinates $z^1, \dots, z^n$, a family of functions $\Gamma^k_{pq}(z)$ which transform under arbitrary coordinate changes $z \to z(z')$ according to the formula

$$\Gamma^{r'}_{i'k'} = \Gamma^r_{ik}\frac{\partial z^i}{\partial z^{i'}}\frac{\partial z^k}{\partial z^{k'}}\frac{\partial z^{s'}}{\partial z^r} + \frac{\partial^2 z^r}{\partial z^{i'}\partial z^{k'}}\frac{\partial z^{s'}}{\partial z^r}. \tag{6.18}$$

The quantities Γ^k_{pq} are called *Christoffel symbols*. We then define the *covariant derivative* of a tensor of type (p,q) by

$$\nabla_r \tilde{T}^{(k)}_{(l)} \equiv \tilde{T}^{(k)}_{(l);r} = \frac{\partial \tilde{T}^{(k)}_{(l)}}{\partial z^r} + \sum_{a=1}^{p}\Gamma^{k_a}_{sr}\tilde{T}^{k_1\dots(k_a\to s)\dots k_p}_{l_1\dots l_q} - \sum_{a=1}^{q}\Gamma^s_{l_a r}\tilde{T}^{k_1\dots k_p}_{l_1\dots(l_a\to s)\dots l_q}. \tag{6.19}$$

Terminology: We often call the operation of covariant differentiation a *connection* while the Christoffel symbols are called the *connection coefficients*.

Definition (Euclidean/Affine Connection): A connection Γ^i_{jk} is *Euclidean or affine* if there exists a basis of coordinates for which $\Gamma^i_{jk} = 0$. The coordinates themselves are often referred to by the same name in this case.

Definition (Torsion Tensor): The Christoffel Symbols are not components of a tensor, however the anti-symmetric combination

$$T^i_{jk} = \Gamma^i_{jk} - \Gamma^i_{kj} = \Gamma^i_{[jk]} \tag{6.20}$$

is a tensor and we call it the *torsion tensor*.

Definition (Symmetric Connection): A connection is said to be *symmetric or torsion-free* if the torsion tensor is identically zero ($\Gamma^i_{jk} = \Gamma^i_{kj}$).

Proposition (Properties of Covariant Differentiation): Aside from the obvious linearity, there are two main pertinent properties of covariant differentiation:

1. **Commutativity with Contraction:** It is sufficient to show this for tensors of type $(1,1)$. Since $T = T^k_k$ is a scalar, we can use the definition

   $$\nabla_r T^k_l = \frac{\partial T^k_l}{\partial z^r} + \Gamma^k_{sr}T^s_l - \Gamma^s_{lr}T^k_s$$

   to evaluate the covariant derivative of the contracted tensor $T = T^k_k$:

   $$\nabla_r T^k_k = \frac{\partial T^k_k}{\partial z^r} + \Gamma^k_{sr}T^s_l - \Gamma^s_{lr}T^k_s = \partial_r T$$

2. **Leibniz Rule for Products:** The covariant derivative of the tensor product follows the Leibniz rule, i.e. if $T^{(i)(j)}_{(p)(q)} = R^{(i)}_{(p)} S^{(j)}_{(q)}$, then the covariant derivative is given by
$$\nabla_k T^{(i)(j)}_{(p)(q)} = (\nabla_k R^{(i)}_{(p)}) S^{(j)}_{(q)} + R^{(i)}_{(p)} (\nabla_k S^{(j)}_{(q)}). \tag{6.21}$$

These properties together with the formulae we have derived for covariant derivatives of scalars and vectors *uniquely determine* the tensor operation of covariant differentiation corresponding to a given connection Γ^i_{jk}.

Thus the next theorem follows immediately.

Theorem (Unique Covariant Derivative Determined by a Connection): Let Γ^i_{jk} be a connection. If a tensor operation satisfies the four conditions:

1. It is linear and commutes with the operation of contraction.
2. On rank-0 tensors the operation coincides with the standard partial derivative $\nabla_k T = \partial_k T$.
3. On vectors the operation is given by
$$\nabla_k T^i = \frac{\partial T^i}{\partial z^k} + \Gamma^i_{rk} T^r.$$
4. On a product $T^{(i)(j)}_{(p)(q)} = R^{(i)}_{(p)} S^{(j)}_{(q)}$ of tensors the operation acts according to the Leibniz rule
$$\nabla_k T^{(i)(j)}_{(p)(q)} = (\nabla_k R^{(i)}_{(p)}) S^{(j)}_{(q)} + R^{(i)}_{(p)} (\nabla_k S^{(j)}_{(q)}).$$

Then it is the operation of covariant differentiation determined by the connection and on general tensors the operation is given by (6.19).

Now we have constructed a derivative that will act on any relevant object (scalars, vectors, tensors) of interest to us in any relevant coordinate system. This machinery is obviously very powerful and to understand it a bit better we will now think of it in a more geometric way.

6.3 Parallel Transport of Fields

In what follows let $\xi \in T_P M$ be a vector in the tangent space of a point P on the m-dimensional manifold M. We consider a general tensor of type (p, q) defined as $T = (T^{(i)}_{(j)})$.

Definition (Directional Derivative of a Tensor): The *directional derivative* of T at P along (or relative to) the vector ξ defined as
$$\nabla_\xi T^{(i)}_{(j)} = \xi^k \nabla_k T^{(i)}_{(j)}, \tag{6.22}$$

is a tensor of the same type (p,q) as T. For a scalar f it coincides with the normal directional derivative (1.20),

$$\nabla_\xi f = \xi^k \partial_k f.$$

Now to visualise the importance of the connection, let $\xi(t)$ be a velocity vector of some curve $\mathcal{C}$ on M in the coordinate system

$$x^i = x^i(t), \qquad \xi^i(t) = \frac{dx^i}{dt}, \qquad i = 1, \dots, m.$$

If $\partial_\xi f = 0$ for the scalar field f for all points of $\mathcal{C}$ then $f(x^1(t), \dots, x^m(t)) = \text{const}$. But in comparison, the question of whether a vector or tensor field is constant along the curve has no significance because the components of a tensor of rank greater than zero are completely coordinate dependent. A given connection gives us a way to compare two tensors attached to different points (with different local coordinates) of the manifold M.

Definition (Covariantly Parallel Tensors): Let Γ^i_{jk} be a connection defined on a manifold M with local coordinates $x^1, \dots, x^m$ in some chart U. Let $x^i(t)$ for $t \in [a,b]$ be a segment $\mathcal{C}$ of an arbitrary piece-wise smooth curve on M. We say that a tensor field T is *covariantly constant or parallel* along $\mathcal{C}$ if

$$\nabla_\xi T = \xi^k \nabla_k T = 0, \qquad \xi^k = \frac{dx^k}{dt} \tag{6.23}$$

is satisfied. For a vector field T^i the explicit formula is

$$\nabla_\xi T^i = \xi^k \nabla_k T^i = \frac{dx^k}{dt}\left(\frac{\partial T^i}{\partial x^k} + \Gamma^i_{jk} T^j\right) = \frac{dT^i(x(t))}{dt} + \frac{dx^k}{dt}\Gamma^i_{jk}T^j. \tag{6.24}$$

The concept of covariant parallelism is only dependent on the connection and the specific curve $\mathcal{C}$, it does not depend on coordinates because covariant differentiation is a tensor operation. Furthermore there is generally no tensor field which would be parallel along all curves in a chart U. The main exception to this is if the connection is Euclidean and $\nabla_\xi T^{(i)}_{(j)} = 0$. In this case we find that $T^{(i)}_{(j)} = \text{const}$ along $\mathcal{C}$ and extending this to all points of U we get a tensor field that is parallel to any other curve in a chart U. Furthermore a product of covariantly parallel tensors is covariantly parallel.

The concept becomes less clear if the connection is not Euclidean, unlike the intuitive 'straight lines that don't meet' definition, we now have a more abstract answer to the question of what it means for two vectors at two distinct points to be parallel. Let's think about this visually on the manifold S^2 in Figure (6.1).

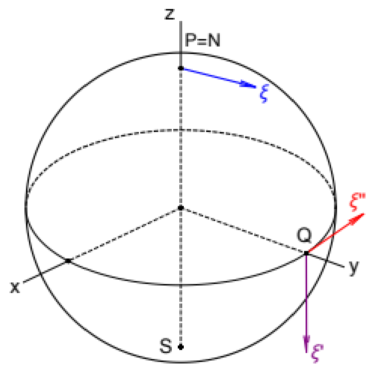


Figure 6.1: Two vectors ξ' and ξ'' in the tangent space $T_Q S^2$ and one vector ξ in the tangent space $T_P S^2$.

To find out if ξ is parallel to ξ' or ξ'' we need to move the vector along the manifold from P to Q. To compare ξ and ξ' we move along the circle in the yz-plane. While traversing the manifold a tangent vector must be kept tangent to the manifold. Clearly at Q the two vectors would be parallel after ξ had been moved in this way.

At first glance the question of whether ξ and ξ'' would seem strange since they are pointing in different directions, however if we take a different path in moving from P to Q, namely the path described below in Figure (6.2).

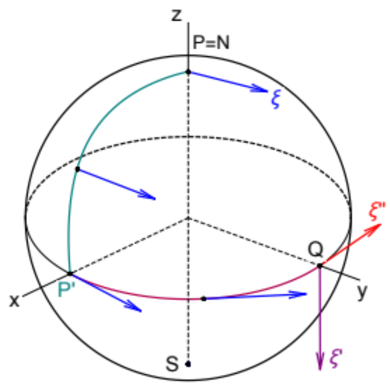


Figure 6.2: The transportation of ξ from P to Q first by moving along the circle in the xz-plane to the point P' and then along the equator until we reach Q.

So we see that depending on the path chosen between P and Q, ξ can be shown to be parallel to both other vectors. This suggests the need for a way of moving one covariantly parallel vector to another one and we call it parallel transport.

Definition (Parallel Transport): Let $(T^i)_P$ be a vector at a point $P(x_0^1, \ldots, x_o^m)$ and let $x^i(t)$, $t \in [0,1]$ be a curve segment $\mathcal{C}$ joining P to a point $Q(x_1^1, \ldots, x_1^m)$. The unique vector field $(T^i)_P$ defined at all points of $\mathcal{C}$, taking the value $(T^i)_P$ at $P(t = 0)$ and parallel along $\mathcal{C}$ is said to result from *parallel transport* of the vector $(T^i)_P$ along $\mathcal{C}$ to $Q(t = 1)$.

The value of (T^i) at Q is denoted by $(T^i)_Q$ and its called the *result of parallel transport* of $(T^i)_P$ along $\mathcal{C}$ to Q relative to the given correlation. The equation which determines the vector field at all points along the curve is the equation of parallel transport (6.24)

$$\frac{dx^k}{dt}\nabla_k T^i = \frac{dx^k}{dt}\left(\frac{\partial T^i}{\partial x^k} + \Gamma^i_{jk}T^j\right) = \frac{dT^i(x(t))}{dt} + \frac{dx^k}{dt}\Gamma^i_{jk}T^j = 0$$

and the initial conditions $T^i(0) = T^i$.

7 Geometry of Connections

7.1 Geodesics

Now we must think about which curves, on an arbitrary manifold (which may or may not have a metric), play the role of a *straight line* between two points. Well in Euclidean geometry, a straight line is distinguished by two properties:

1. It is the shortest curve between two points (need a metric to define *shortest*).
2. Velocity vectors at all points of the curve are parallel.

Obviously for the wildly varying topologies we find on a manifold, the second property will have to do.

Definition (Geodesics): A curve $x^i = x^i(t)$ is called a *geodesic* if the vector field defined by its tangent vector $T^i = \frac{dx^i}{dt}$ is parallel along the curve itself. In other words, if the curve parallel transports its own tangent vector

$$\nabla_T(T) = \nabla_{\dot{x}}(\dot{x}) = 0. \tag{7.1}$$

We can also write this in components which yields

$$\nabla_T(T)^i = \frac{dx^k}{dt}\nabla_k\left(\frac{dx^i}{dt}\right) = \frac{d}{dt}\frac{dx^i}{dt} + \frac{dx^k}{dt}\Gamma^i_{jk}\frac{dx^j}{dt} = 0. \tag{7.2}$$

This gives us that geodesics are defined by the equation

$$\frac{d^2x^k}{dt^2} + \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0. \tag{7.3}$$

If the connection is Euclidean then the solutions are straight lines in the normal sense. The geodesics also do not depend on the torsion of the connection, only the symmetric part $\Gamma^i_{(jk)} = \Gamma^i_{jk} + \Gamma kj^i$. And the following theorem will give us the fact that solutions to this second order ODE are unique.

Theorem (Uniqueness of Geodesics): Let Γ^i_{kj} be a connection defined on a manifold M. Then for any point P in some chart U and any vector $(T^i)_P$ attached to the point, there exists a unique geodesic starting from P with the initial tangent vector $(T^i)_P$.

7.2 Connections Compatible with the Metric

We are able to define geodesics without the notion of distance on a manifold because the connection and the metric are completely independent. However, for manifolds which have a metric we do have a notion of distance and it would be nice if we could relate the metric to the connection. Then we could define the set of geodesics connecting two nearby points to contain the *shortest curve*. Furthermore, a metric can be used to lower the indices of a tensor and going forward we want the covariant derivative to commute with this operation. Lastly we would like to use the connection to parallel transport the metric defined at a point P to any other point of M in a path-independent way.

Definition (Connection Compatible with the Metric): A connection Γ^i_{kj} is said to be *compatible with the metric* g_{ij} if the covariant derivative of the metric tensor is identically zero

$$\nabla_k g_{ij} \equiv 0, \qquad i, j, k = 1, \ldots, m \tag{7.4}$$

Relating this definition to the other concepts we have introduced, we can say that if it is satisfied for Γ^i_{kj} and g_{ij} then the metric tensor is *covariantly constant / parallel* along any curve $\mathcal{C}$. Hence, the result of parallel transport of (g_{ij}) is independent of the curves used.

Proposition (Lowering Commutes with Covariant Differentiation): Given a connection compatible with the metric, the associated covariant derivative operation commutes with the operation of lowering any index of a tensor

$$\nabla_k (g_{lm} T^{(i)}_{(j)}) = (\nabla_k g_{lm}) T^{(i)}_{(j)} + g_{lm} (\nabla_k T^{(i)}_{(j)}) = g_{lm} (\nabla_k T^{(i)}_{(j)}). \tag{7.5}$$

When we have a metric we also have the bilinear operation of a scalar product between vector fields on the manifold and this leads to the following property for a connection compatible with the metric.

Proposition (Product of Parallel Vector Fields is Parallel): If vector fields $T^i(t)$ and $S^i(t)$ are both parallel along a curve $x^i = x^i(t)$, then their *scalar product is covariantly constant / parallel* along the curve:

$$\frac{d}{dt}\langle T, S\rangle = \frac{d}{dt}(g_{ij} T^i S^j) = \frac{dx^k}{dt} \nabla_k (g_{ij} T^i S^j) = g_{ij} \frac{dx^k}{dt} \nabla_k (T^i S^j) = 0. \tag{7.6}$$

As a technical remark, it can be said that parallel transport of vectors from a point P to a point Q along a given curve defines an *orthogonal transformation* from the tangent space at P to the tangent space at Q.

We now consider the special case when the connection compatible with the metric is symmetric or *torsion-free*. It turns out that we can always find such a connection when the metric is non-singular by the following theorem.

Theorem (Existence of a Unique, Symmetric and Compatible Connection): If the metric g_{ij} is non-singular on a chart U of the manifold M being considered, then there is a *unique, torsion-free* (symmetric) connection compatible with the metric. It is given in any system of coordinates $x^1, \ldots, x^m$ by the so-called Christoffel's formula

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl} \bigg(\frac{\partial g_{lj}}{\partial x^i} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \bigg). \tag{7.7}$$

Proof: First we lower the indices of the connection

$$\Gamma_{j,ik} \equiv g_{pj} \Gamma^p_{ik} = g_{jp} \Gamma^p_{ik}.$$

Then if we expand the covariant derivative of the metric tensor (which must vanish identically) we get

$$\nabla_k g_{ij} = \partial_k g_{ij} - g_{pj} \Gamma^p_{ik} - g_{ip} \Gamma^p_{jk} = \partial_k g_{ij} - \Gamma_{j,ik} - \Gamma_{i,jk} = 0$$

The last expression also vanishes when we cyclically permute the indices so we produce an additional two identities

$$\partial_j g_{ki} - \Gamma_{i,kj} - \Gamma_{k,ij} = 0$$

$$\partial_i g_{jk} - \Gamma_{k,ji} - \Gamma_{j,ki} = 0$$

If we sum the last two equations and and subtract the result from the first we obtain

$$\partial_j g_{ki} + \partial_i g_{jk} - \partial_k g_{ij} - 2\Gamma_{k,ij} = 0 \tag{7.8}$$

where we have used that $\Gamma_{k,ij} = \Gamma_{k,ji}$. Then solving for the connection we get

$$\Gamma_{k,ij} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}) \implies \Gamma^k_{ij} = \frac{1}{2} g^{kp}(\partial_i g_{pj} + \partial_j g_{pi} - \partial_p g_{ij}) \tag{7.9}$$

One additional property for symmetric and compatible (with the metric) connections is that if $\partial_i g_{jk} = 0$ at a given point for any value of the indices, then at that point, all symbols vanish $\Gamma^k_{ij} = 0$. The theorem also leads to the following result for the divergence of a vector field.

Proposition (Stokes from the Covariant Derivative): The divergence of a vector field in terms of the connection is

$$\mathrm{div} T = \nabla_i T^i = \frac{\partial T^i}{\partial x^i} + \Gamma^i_{ki} T^k.$$

Which we know from (5.33) can be expressed as

$$\nabla_i T^i = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^i} (\sqrt{|g|} T^i).$$

Then by using the contracted version of Christoffel's formula

$$\Gamma^i_{ki} = \frac{1}{2} g^{ip}(\partial_k g_{pi} + \partial_i g_{kp} - \partial_p g_{ki}) = \frac{1}{2} g^{ip} \partial_k g_{pi} = \frac{1}{2g} \partial_k g$$

we get

$$\sqrt{|g|}\nabla_i T^i = \frac{\partial}{\partial x^i}(\sqrt{|g|}T^i) \tag{7.10}$$

and therefore if M is an oriented and closed manifold (an oriented compact manifold without boundary), we obtain

$$\int_M \nabla_i T^i \Omega_M = \int_M \nabla_i T^i \sqrt{|g|} dx^1 \wedge \cdots \wedge dx^m = 0. \tag{7.11}$$

This follows from Stokes' theorem for manifolds without boundary.

7.3 Geodesics and the Metric

Now that we have related the metric to a compatible connection in a unique way using Christoffel's formula, we can then redefine the set of geodesics to be the set of curves connecting two nearby points on the manifold to be the shortest ones. As we will see this does not interfere with the definition of curves which parallel transport their own tangent vector.

The length of a curve $\mathcal{C} : x^i = x^i(t)$, $t \in [a,b]$ between the points P and Q on a manifold which has a connection compatible with the metric is given by

$$L = \int_{\mathcal{C}} ds = \int_a^b \sqrt{g_{ij}\dot{x}^i\dot{x}^j} dt = \int_a^b \mathcal{L}(x,\dot{x}) dt. \tag{7.12}$$

The length L is minimised by the corresponding solutions of the variational principle which are found by setting $\delta L = 0$. This yields

$$\delta L = \int_a^b \delta\mathcal{L} dt \int_a^b \left(\frac{\partial \mathcal{L}}{\partial x^k}\delta x^k + \frac{\partial \mathcal{L}}{\partial \dot{x}^k}\delta \dot{x}^k \right) dt \tag{7.13}$$

$$\overset{IBP}{=} \int_a^b \frac{\partial \mathcal{L}}{\partial x^k}\delta x^k - \int_a^b \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}^k}\delta x^k = 0.$$

Here, we used that the arbitrary variations must vanish at the endpoints since the endpoints of the curve are fixed. Furthermore since δx^k is an arbitrary deformation, we must have that its coefficient terms vanish to make the equality correct. This gives us the familiar Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}^k} - \frac{\partial \mathcal{L}}{\partial x^k} = 0. \tag{7.14}$$

Computing the relevant derivatives of the line element gives

$$\frac{\partial \mathcal{L}}{\partial \dot{x}^k} = \frac{1}{\mathcal{L}} g_{lk}\dot{x}^l, \qquad \frac{\partial \mathcal{L}}{\partial x^k} = \frac{1}{2\mathcal{L}} \partial_k g_{lj}\dot{x}^l\dot{x}^j,$$

and

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}^k} = \frac{1}{\mathcal{L}} g_{lk}\ddot{x}^l + \frac{1}{\mathcal{L}}\partial_j g_{lk}\dot{x}^l\dot{x}^j - \frac{\dot{\mathcal{L}}}{\mathcal{L}^2} g_{lk}\dot{x}^l.$$

This gives us that the Euler-Lagrange equations are

$$g_{lk}\ddot{x}^l + \partial_j g_{lk}\dot{x}^l\dot{x}^j - \frac{1}{2}\partial_k g_{lj}\dot{x}^l\dot{x}^j = \frac{\dot{\mathcal{L}}}{\mathcal{L}} g_{lk}\dot{x}^l$$

where we have multiplied across by $\mathcal{L}$. Now we multiply the whole equation with the metric tensor g^{ik} which gives

$$\ddot{x}^i + g^{ik}\partial_j g_{lk}\dot{x}^l\dot{x}^j - \frac{1}{2} g^{ik}\partial_k g_{lj}\dot{x}^l\dot{x}^j = \frac{\dot{\mathcal{L}}}{\mathcal{L}}\dot{x}^i$$

$$\Longrightarrow \ddot{x}^i + \frac{1}{2} g^{ik}(\partial_j g_{lk} + \partial_l g_{jk} - \partial_k g_{lj})\dot{x}^l\dot{x}^j = \frac{\dot{\mathcal{L}}}{\mathcal{L}}\dot{x}^i.$$

We recognise the presence of the unique connection compatible with the metric from the term in brackets so our equation reduces to

$$\ddot{x}^i + \Gamma^i_{lj}\dot{x}^l\dot{x}^j = \frac{1}{2}\dot{x}^i \frac{d}{dt}\ln(g_{kl}\dot{x}^k\dot{x}^l). \tag{7.15}$$

From this equation we can retrieve the geodesic equation by demanding the right hand side vanishes. This is equivalent to the condition:

$$g_{ij}\dot{x}^i\dot{x}^j \equiv K = \text{const.} \tag{7.16}$$

Geometrically, the quantity K is the square of the length of the tangent vector to the curve. For any given parameterisation of the curve $\mathcal{C}$ with parameter t, one can find a new one parameterised by $s = s(t)$ such that K is a constant. This (more pertinent) parameterisation is given by solving the equation

$$K = g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\left(\frac{dt}{ds}\right)^2 \Longrightarrow \frac{ds}{dt} = \sqrt{\frac{g_{ij}\dot{x}^i\dot{x}^j}{K}} \tag{7.17}$$

which has the solution

$$s(t) = \int_a^t \sqrt{\frac{g_{ij}\dot{x}^i\dot{x}^j}{K}}\, d\tau. \tag{7.18}$$

This solution for $s(t)$ is a generalisation of the arc length parameterisation called the *affine parameterisation.* In this parameterisation, the length of the tangent vectors remain constant while being parallel transported along the geodesic.

8 Curvature Tensors

In this section we will put a rigorous framework around the concept of curvature. We will come to find that the curvature is quantified by the Riemann tensor, which is derived from the connection. The motivation for such a measure of curvature is to do with what we know about the "flatness" of a connection associated with a Euclidean metric. The conventional connection associated with the Euclidean metric is as we know given by Christoffel's formula and it has various properties which can be thought of as manifestations of flatness. For example the parallel transport around a closed loop leaves a vector unchanged, covariant derivatives commute and that initially parallel geodesics remain parallel. In our example of the parallel transport of a vector ξ from P to Q on S^2, we saw that the vector actually transformed rather than remaining unchanged. So the tensor that we wish to derive is the one which measures the curvature enclosed by the closed loop since curvature is what determines the transformation of the vector.

8.1 Derivation of the Riemann Tensor

Using what we already know we can make an estimate for the definition of the curvature tensor. The changed experienced by a vector when it is parallel transported in the most general case will depend on the curvature tensor R acting on the vector T^μ itself and two more vectors A^ν and B^ρ which define the closed loop. The expression should be anti-symmetric in the ν, ρ indices since swapping the order of A and B corresponds to traversing the loop in the opposite direction. There will also be an upper and lower index on R to reflect the fact it is a matrix multiplying a vector. Clearly we should obtain an expression of the form

$$\delta T^\mu = R^\mu_{\sigma\rho\nu} T^\sigma A^\rho B^\nu. \tag{8.1}$$

The curvature tensor $R^\mu_{\sigma\rho\nu}$ that we obtain here is not fully defined since we still don't have a reason to order the indices in this way. So while informative, we will have to derive the tensor properly using another method.

The easiest way to derive the curvature tensor is to actually consider the commutator of covariant derivatives. This is directly related to the parallel transport around a loop because the covariant derivative of a tensor in a direction along which it is parallel is zero by definition. So the commutator of two covariant derivatives measures the difference between parallel transporting the tensor first one way and then the other versus the opposite ordering. Consider any arbitrary connection Γ^i_{jk}, and a vector field T^i, then we compute the commutator of two covariant derivatives as

$$\begin{aligned}
[\nabla_k, \nabla_l]\xi^i &= \partial_k(\partial_l \xi^i + \xi^p \Gamma^i_{pl}) + \nabla_l \xi^p \Gamma^i_{pk} - \nabla_q \xi^i \Gamma^q_{lk} - (k \longleftrightarrow l) \\
&= \partial_k \xi^p \Gamma^i_{pl} + \xi^p \partial_k \Gamma^i_{pl} + \nabla_l \xi^p \Gamma^i_{pk} - \nabla_q \xi^i \Gamma^q_{lk} - (k \longleftrightarrow l) \\
&= (\nabla_k \xi^p - \xi^q \Gamma^p_{qk}) \Gamma^i_{pl} + \xi^p \partial_k \Gamma^i_{pl} + \nabla_l \xi^p \Gamma^i_{pk} - \nabla_q \xi^i \Gamma^q_{lk} - (k \longleftrightarrow l)
\end{aligned}$$

$$= -\xi^q \Gamma^p_{qk} \Gamma^i_{pl} + \xi^p \partial_k \Gamma^i_{pl} + \left(\nabla_k \xi^p \Gamma^i_{pl} + \nabla_l \xi^p \Gamma^i_{pk}\right) - \nabla_q \xi^i \Gamma^q_{lk} - (k \longleftrightarrow l)$$

The term in brackets cancels under the anti-symmetrisation so we are left with

$$-\xi^q \Gamma^p_{qk} \Gamma^i_{pl} + \xi^p \partial_k \Gamma^i_{pl} - \nabla_q \xi^i \Gamma^q_{lk} - (k \longleftrightarrow l)$$

$$= \partial_k \Gamma^i_{jl} \xi^j - \Gamma^p_{jk} \Gamma^i_{pl} \xi^j - \nabla_j \xi^i \Gamma^j_{lk} - (k \longleftrightarrow l)$$

$$= (\partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} + \Gamma^p_{jl} \Gamma^i_{pk} - \Gamma^p_{jk} \Gamma^i_{pl}) \xi^j + (\Gamma^j_{kl} - \Gamma^j_{lk}) \nabla_j \xi^i. \tag{8.2}$$

So what have we discovered in this expression. The last term involving a difference of connections is clearly just the torsion tensor T^j_{kl}. The left hand side of the equation is of course a tensor from our definition of covariant derivatives. This leaves the fact that the first term in brackets multiplying the vector field ξ^j must also be a tensor and so we identify it as the Riemann (curvature) tensor R^i_{jkl} which we can now define as

$$R^i_{jkl} = \partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} + \Gamma^p_{jl} \Gamma^i_{pk} - \Gamma^p_{jk} \Gamma^i_{pl} \tag{8.3}$$

The commutator of covariant derivatives can now be written in a more compact form known as the **Ricci identity**

$$[\nabla_k, \nabla_l] \xi^i = R^i_{jkl} \xi^j + T^j_{kl} \nabla_j \xi^i. \tag{8.4}$$

Now there are several non-trivial elements implicit in the derivation we have completed here. Firstly it appears that the Riemann tensor measures the part of the commutator that is proportional to the vector field while the torsion tensor measures the part that is proportional to the covariant derivative of the vector field and the second derivative in the commutator does not enter into the expression independently at all. We also observe that the hypothesised anti-symmetry of R in its last two indices has emerged in the expression (8.3). Lastly, we have constructed the curvature tensor completely from the connection, with no mention of any metric. This ensured that the curvature tensor is as general as possible for the case when the connection is not compatible or symmetric. One good thing to check would be that the curvature tensor vanishes when we have an affine/Euclidean connection.

Proposition (The Curvature Tensor for an Affine Connection): Given a connection Γ^i_{jk}, assume there exist a set of affine coordinates such that $\Gamma^i_{lk} = 0$. Then in these coordinates the covariant derivative of a vector field ξ^j is simply the usual partial derivative and the commutator of covariant derivatives is trivial

$$[\nabla_k, \nabla_l] \xi^j = (\partial_k \partial_l - \partial_l \partial_k) \xi^j = 0 \tag{8.5}$$

by the *equality of mixed partial derivatives.* Then for this connection the curvature tensor and torsion tensor identically vanish

$$R^i_{jkl} = 0, \qquad T^j_{kl} = 0. \tag{8.6}$$

Since the left hand side of (8.5) is a tensor, this holds in all other coordinate systems.

We can reformulate the curvature tensor and the torsion tensor by thinking of them as multilinear maps of the vector fields. Thinking of the torsion tensor as a map from two vector fields ξ, η to a third ζ we write in components

$$(T(\xi,\eta))^i = T^i_{kl}\xi^k\eta^l. \tag{8.7}$$

In a similar way we define the corresponding map of the curvature tensor from three vector fields to a fourth one in components as

$$(R(\xi,\eta)\zeta)^i = R^i_{jkl}\xi^k\eta^l\zeta^j. \tag{8.8}$$

With these definitions we can write the coordinate-free formulation of the torsion and curvature tensors which we will prove in the following lemma.

Lemma (Coordinate-Free Torsion and Curvature Tensors): For arbitrary vector fields ξ, η, ζ, the following equations hold

$$T(\xi,\eta) = \nabla_\eta\xi - \nabla_\xi\eta + [\xi,\eta], \tag{8.9}$$

$$R(\xi,\eta)\zeta = [\nabla_\xi, \nabla_\eta]\zeta - \nabla_{[\xi,\eta]}\zeta. \tag{8.10}$$

Proof: We first show that these equations are linear in the vectors. If we replace $\xi \to f\xi$ where f is a function and use the fact that

$$\nabla_{f\xi}\eta = f\nabla_\xi\eta, \qquad [f\xi,\eta] = f[\xi,\eta] - (\partial_\eta f)\xi.$$

Then we obtain

$$T(f\xi,\eta) = \nabla_\eta(f\xi) - \nabla_{f\xi}\eta + [f\xi,\eta]$$
$$= f(\nabla_\eta\xi - \nabla_\xi\eta + [\xi,\eta]) + (\nabla_\eta)f\xi - (\partial_\eta f)\xi = fT(\xi,\eta)$$

where the last two terms on the left hand side cancel each other. Now for the curvature tensor we obtain

$$R(f\xi,\eta) = (\nabla_{f\xi}\nabla_\eta - \nabla_\eta\nabla_{f\xi})\zeta - \nabla_{[f\xi,\eta]}\zeta$$
$$= f\nabla_\xi\nabla_\eta\zeta - \nabla_\eta(f\nabla_\xi\zeta) - f\nabla_{[\xi,\eta]}\xi + \partial_\eta f\nabla_\xi\zeta$$
$$= f([\nabla_\xi,\nabla_\eta] - \nabla_{[\xi,\eta]})\zeta - \nabla_\eta f\nabla_\xi\zeta + \partial_\eta f\nabla_\xi\zeta = fR(\xi,\eta)\zeta$$

where the last two terms on the left hand side cancel once again. The linearity in η follows in the exact same way but now we check that the equations are linear in ζ by setting $\zeta \to f\zeta$ and applying a rather large product rule

$$R(\xi,\eta)(f\zeta) = f([\nabla_\xi,\nabla_\eta]\zeta - \nabla_{[\xi,\eta]}\zeta) + \nabla_\eta f\nabla_\xi\zeta + \nabla_\xi f\nabla_\eta\zeta$$

$$-\nabla_\xi f\nabla_\eta\zeta - \nabla_\eta f\nabla_\xi\zeta + ((\nabla_\xi\nabla_\eta - \nabla_\eta\nabla_\xi)f)\zeta - (\nabla_{[\xi,\eta]}f)\zeta$$
$$= f([\nabla_\xi,\nabla_\eta]\zeta - \nabla_{[\xi,\eta]}\zeta) = fR(\xi,\eta)\zeta.$$

With linearity in the vector fields proven, it is sufficient to show that the equations are satisfied for the basis vector fields $(\xi, \eta, \zeta) = (e_k, e_l, e_j)$ which are oriented along the corresponding coordinate axis of their index. We will use here that for unit normal basis vectors along a coordinate axis, the commutator $[e_i, e_j] = [\partial_i, \partial_j] = 0$ vanishes by the equality of mixed partials. For the torsion tensor we have in components

$$(\nabla_{e_l} e_k - \nabla_{e_k} e_l + [e_k, e_l])^i = (\nabla_l e_k - \nabla_k e_l)^i = \Gamma^i_{kl} - \Gamma^i_{lk} = T^i_{kl},$$

Where we have used that by definition $\nabla_i e_j = \Gamma^p_{ji} e_p$. Similarly we check the equation holds for the curvature tensor which yields

$$([\nabla_{e_k}, \nabla_{e_l}]e_j - \nabla_{[e_k,e_l]} e_j)^i = ([\nabla_k, \nabla_l]e_j)^i$$

$$= (\nabla_k(\Gamma^p_{jl} e_p) - \nabla_l(\Gamma^p_{jk} e_p) = \partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} + \Gamma^p_{jl}\Gamma^i_{pk} - \Gamma^p_{jk}\Gamma^i_{pl} = (R(e_k, e_l)e_j)^i.$$

This concludes the proof □.

8.2 Symmetries of The Curvature Tensor

A natural question to propose about the general Riemann tensor we have derived is about the number of independent dimensions it has. The naive answer is n^4 but it has various properties which reduce the number of independent components especially when we specialise to the compatible case with a given metric. These properties are categorised in the following theorem.

Theorem (Symmetries of R^i_{jkl}): The Riemann tensor satisfies the following identities, each dependent on the given connection Γ^i_{jk}:

1. **Any Γ^i_{jk}:** The curvature tensor is antisymmetric in its last two indices

$$R^i_{jkl} + R^i_{jlk} = 0. \tag{8.11}$$

2. **Torsion-Free Γ^i_{jk}:** The sum of cyclic permutations in the last three indices and therefore the anti-symmetric part of the last three indices vanishes

$$R^i_{jkl} + R^i_{ljk} + R^i_{klj} = 0 \iff R^i_{[jkl]} = 0 \tag{8.12}$$

3. **Compatible Γ^i_{jk}:** Then the connection is given by the Christoffel formula and lowering the upper index of the curvature tensor using the metric $R_{ijkl} = g_{ip} R^p_{jkl}$, we find that the tensor is anti-symmetric in the first two indices

$$R_{ijkl} + R_{jikl} = 0. \tag{8.13}$$

4. **Torsion-Free and Compatible Γ^i_{jk}:** The curvature tensor is symmetric under the exchange of its first pair and last pair of indices $(ij) \longleftrightarrow (kl)$

$$R_{ijkl} - R_{klij} = 0. \tag{8.14}$$

Lemma (Existence of Normal Coordinates): Given a point p on a manifold M with a chart (x^i), one can find a new chart $(\hat{x}^i)$ such that the transformed connection has a vanishing symmetric contribution at p

$$\hat{\Gamma}^i_{(jk)}(p) \equiv \frac{1}{2}(\hat{\Gamma}^i_{jk}(p) + \hat{\Gamma}^i_{kj}(p)) = 0. \tag{8.15}$$

Which can be rephrased as $\hat{\Gamma}^i_{jk} = \frac{1}{2}\hat{T}^i_{jk}$. We call the set of coordinates $(\hat{x}^i)$, the normal coordinates.

Proof: Let p be the point with coordinates $x^i = 0$, $\hat{x}^i = 0$ in each system. We wish to look for the hatted coordinates such that they are given by

$$\hat{x}^i = x^i + \frac{1}{2}Q^i_{jk}x^j x^k,$$

where $Q^i_{jk} = Q^i_{kj}$ are some constants to be found from our initial condition (8.15). We now claim that for small $|x| = \max_i \hat{x}^i$ we can invert this coordinate transformation such that

$$x^i = \hat{x}^i - \frac{1}{2}Q^i_{jk}\hat{x}^j\hat{x}^k + \mathcal{O}(|x|^3).$$

To find how the connection transforms between the two systems we must compute the first and second derivatives

$$\begin{aligned}
\frac{\partial \hat{x}^i}{\partial x^j} &= \delta^i_j + Q^i_{jk}x^k \qquad \text{which at } p \text{ equals} \qquad \delta^i_j \\
\frac{\partial x^i}{\partial \hat{x}^j} &= \delta^i_j - Q^i_{jk}\hat{x}^k \qquad \text{which at } p \text{ equals} \qquad \delta^i_j \\
\frac{\partial^2 x^i}{\partial \hat{x}^j \partial \hat{x}^k} &= -Q^i_{jk} + \mathcal{O}(|x|) \qquad \text{which at } p \text{ equals} \qquad -Q^i_{jk}.
\end{aligned}$$

According to equation (6.18) we have that the connection transforms between these coordinates at p as

$$\hat{\Gamma}^i_{jk}(p) = \Gamma^i_{jk}(p) - Q^i_{jk}.$$

To get our result all we have to do is choose $Q^i_{jk} = \Gamma^i_{(jk)}(p)$ so that the symmetric part of $\hat{\Gamma}^i_{jk}$ vanishes. □

Remark (Proving The Bianchi Identity Using Normal Coordinates): The Bianchi identities for the curvature tensor of a symmetric connection are

$$\nabla_m R^n_{ikl} + \nabla_l R^n_{imk} + \nabla_k R^n_{ilm} = 0. \tag{8.16}$$

We can always make the symmetric part of a connection vanish at a point p by choosing the normal coordinates which we now denote (x^i), $x^i = 0$. Therefore with these coordinates we have a Euclidean connection $\Gamma^i_{jk} = 0$. At the point p, with these coordinates chosen, the vanishing connection reduces the curvature tensor to

$$R^n_{ikl}(p) = \partial_k \Gamma^n_{il}(p) - \partial_l \Gamma^n_{ik}(p).$$

And the corresponding covariant derivatives in the Bianchi identities to

$$\begin{aligned}
\nabla_m R^n_{ikl} &= \partial_m \partial_k \Gamma^n_{il} - \partial_m \partial_l \Gamma^n_{ik} \\
\nabla_l R^n_{imk} &= \partial_l \partial_m \Gamma^n_{ik} - \partial_l \partial_k \Gamma^n_{im} \\
\nabla_k R^n_{ilm} &= \partial_k \partial_l \Gamma^n_{im} - \partial_k \partial_m \Gamma^n_{il}.
\end{aligned}$$

The sum over these equations vanishes by the equality of mixed partial derivatives proving the identities.

8.3 A Few Important Tensors for GR

Definition (Ricci Tensor): The contraction

$$R_{jl} = R^i_{jil} \tag{8.17}$$

of the Riemann curvature tensor is called the *Ricci tensor*. It is then given explicitly by the equation

$$R_j l = \partial_i \Gamma^i_{jl} - \partial_l \Gamma^i_{ji} + \Gamma^p_{jl} \Gamma^i_{pi} - \Gamma^p_{ji} \Gamma^i_{pl}. \tag{8.18}$$

If the connection is torsion-free and compatible with the metric then $R_{jl} = R_{lj}$.

Now when we make the following two definitions we are mainly referring to Riemannian manifolds with symmetric, compatible connection given by the Christoffel connection.

Definition (Scalar Curvature): The scalar given by the contraction of the Ricci tensor with the metric tensor g^{lj}

$$R = g^{jl} R_{jl} = g^{jl} R^i_{jil} \tag{8.19}$$

is called the scalar curvature.

Definition (Einstein Tensor): If our connection is symmetric and compatible with the metric, then the tensor

$$G_{ij} = R_{ij} - \frac{1}{2} R g_{ij} \tag{8.20}$$

is called the Einstein tensor.

Definition (The Weyl Tensor): The tensor

$$C_{ijkl} = R_{ijkl} - \frac{2}{n-2}\left(g_{i[k}R_{l]j} - g_{j[k}R_{l]i}\right) + \frac{2}{(n-2)(n-1)} g_{i[k}g_{l]j}R \tag{8.21}$$

is called the *Weyl tensor*. The Ricci tensor and scalar contain all of the information about the traces of the Riemann curvature tensor while the Weyl tensor captures all of the traceless parts. The formula is designed in this messy way so that *all possible contractions vanish* while it retains the same symmetries as the Riemann tensor

$$C_{ijkl} = C_{[ij][kl]}, \qquad C_{ijkl} = C_{klij}, \qquad C_{i[jkl]} = 0. \tag{8.22}$$