Chapter 1

Differential Geometry

This chapter will give a rigorous mathematical introduction to differential geometry to lay the groundwork for general relativity. This includes notes on notation.

1.1 Topological Spaces

Topology is the branch of mathematics concerned with extending concepts that are familiar in $\mathbb{R}^n$, such as convergence and continuity, to arbitrary sets. If X is a nonempty set, then a *topology*, $\mathscr{T}$, on X is a collection of subsets of X with the following properties:

1. $X \in \mathscr{T}$ and $\emptyset \in \mathscr{T}$,

2. If $\{U_i\}_{i \in I} \subseteq \mathscr{T}$ for an index set I, then $\bigcup_{i \in I} U_i \in \mathscr{T}$,

3. $T_1 \cap T_2 \in \mathscr{T}$ whenever $T_1, T_2 \in \mathscr{T}$.

That is, a topology $\mathscr{T}$ is a set of subsets of X that is closed under arbitrary unions and finite intersections. The pair $(X, \mathscr{T})$ is called a *topological space*. Often, when the intention is obvious, the topological space is referred to simply as X. The subsets $T \in \mathscr{T}$ are called the *open sets*, and their complements $X \setminus T$, $T \in \mathscr{T}$ are the *closed sets*.

An *open cover* of a subset $S \subseteq X$ is a collection of open sets $\{U_i\}_{i \in I}$ such that $S \subseteq \bigcup_{i \in I} U_i$. The set S is *compact* if every open cover of S has a finite subcover.

An example of a topological space is a *metric space*; a pair (X, d) where X is a set and $d : X \times X \to [0, \infty)$ is a function such that for all $x, y, z \in X$

1. $d(x, y) = 0 \Leftrightarrow x = y$,

2. $d(x, y) = d(y, x)$,

3. $d(x, y) \leq d(x, z) + d(z, y)$.

In a metric space, an *open ball* centred at a with radius $r \in \mathbb{R}^+$ is the set

$$B(a, r) = \{x \in X | d(a, x) < r\}.$$

A subset of X is called *open* if it is a union of open balls. In a metric space (X, d), the *metric topology*, $\mathcal{T}_m$, is the set of open subsets of X with respect to the metric d.

If X is a topological space, and $x_0 \in X$, then a *neighbourhood* of x_0 is any open subset of X containing x_0. A topological space is called *Hausdorff* if any pair of distinct points have disjoint neighbourhoods. Any metric space (X, d) is a Hausdorff space: for any $x, y \in X$, simply take the neighbourhoods to be the open balls $d(x, r)$ and $d(y, r)$ with $r = \frac{1}{2}d(x, y)$.

A space X is *second countable* if there is a countable set $\mathcal{U}$ of open sets in X such that every open set in X is a union of sets in $\mathcal{U}$.

A mapping f between topological spaces X and Y is *continuous* at $x_0 \in X$ if for all open neighbourhoods U of $f(x_0) \in Y$, the preimage $f^{-1}(U)$ contains an open neighbourhood of x_0. It is said to be *continuous* on X if it is continuous at each point.

A *homeomorphism* between X and Y is a continuous *bijection* f (i.e. f is 1-1 and onto) with a continuous inverse f^{-1}. If such an f exists, X and Y are said to be *homeomorphic*.

1.2   Manifolds

Analysis is a familiar concept in $\mathbb{R}^n$, however on general topological spaces it is not applicable in the same way. A manifold is a structure used to describe and investigate certain topological spaces by identifying them with Euclidean space on "small enough" scales. We shall only define and discuss *smooth* manifolds. To make this idea precise and to use it, we require the following new concepts.

Let M be a topological space. A homeomorphism $x : U \to U'$, for open subsets $U \subseteq M$ and $U' \subseteq \mathbb{R}^n$, is called a *coordinate chart*. We write $x = (x^1(p), x^2(p), ..., x^n(p))$, where the x^i, called *(local) coordinates*, are functions of the point $p \in U$. Strictly speaking, the coordinate chart is a map, but commonly it is used to identify the set U with U' and $p \in U$ with $x(p)$ in $\mathbb{R}^n$.

Usually more than one chart is required to cover the whole space, and these charts must be consistent where they overlap. We say two charts (U, x) and (V, y) are *compatible* if $x \circ y^{-1} : y(U \cap V) \to x(U \cap V)$ is smooth.

An *atlas* $\mathcal{A}$ on M is a set of coordinate charts on M such that

1. the domains of the charts, U, cover M

2. any two charts are compatible.

If the compatibility criterion above is applied in the reverse order, it is seen that $y \circ x^{-1} = (x \circ y^{-1})^{-1}$ must also be smooth, that is, the maps between the two charts are smooth homeomorphisms with smooth inverses. Such maps are called *diffeomorphisms*. The set of all charts compatible with every chart in an atlas $\mathcal{A}$ is called the *maximal atlas* containing $\mathcal{A}$ and is unique. A maximal atlas may also be called a *smooth structure* on M.

A *(smooth) manifold* M is a second countable Hausdorff space equipped with a maximal atlas $\mathcal{A}$. If the charts in $\mathcal{A}$ map $U \in M \to \mathbb{R}^n$, then n is called the *dimension* of the manifold. Sometimes M is referred to as an *n-manifold*. If N is a closed subset of an n-manifold M, we say that N is a *k-submanifold* of M if N can be covered by charts (x, U) in M such that

$x(N \cap U) \subset \mathbb{R}^k \times \mathbf{0}$, where $\mathbf{0}$ is the zero of $\mathbb{R}^{n-k}$.

The *Jacobian* of a map $y : \mathbb{R}^n \to \mathbb{R}^n$, $y(x^1, \ldots, x^n) = (y^1, \ldots, y^n)$ is defined to be the matrix

$$J_{ij} = \frac{\partial y^i}{\partial x^j}.$$

On a manifold M, two charts (U, x) and (V, y) are said to have the same *orientation* if the diffeomorphism

$$x \circ y^{-1} : y(U \cap V) \to x(U \cap V)$$

has positive Jacobian determinant. An atlas $\mathcal{A}$ is *oriented* if each pair of overlapping charts have the same orientation. A manifold M is *orientable* if it has an oriented atlas.

Let M be a smooth manifold with an atlas $\mathcal{A} = (x_\alpha, U_\alpha)_{\alpha \in I}$ for some index set I. A *partition of unity subordinate to* (U_α) is a family of smooth functions $(\psi_\beta)_{\beta \in J}$, $\psi_\beta : M \to [0,1]$ with compact support such that:

1. for each $\beta \in J$ there is $\alpha \in I$ such that $\mathrm{supp}\psi_\beta \subset U_\alpha$,

2. for every compact subset $S \in M$, $S \cup \mathrm{supp}\psi_\beta \neq \emptyset$ for only finitely many β,

3. $\sum_{\beta \in J} \psi_\beta(x) = 1$ for all $x \in M$.

Theorem 1.2.1 *Let M be a smooth manifold, and let (x_α, U_α) be an atlas on M. Then there exists a partition of unity subordinate to (U_α).*

For any open cover (U_α) we can find a refinement such that we can use the same index set for the partition of unity as we use for the cover. For a proof of this statement and Theorem 1.2.1 see [?].

Examples of manifolds include:

1. $\mathbb{R}^n$ itself, which is covered by a single chart, namely the identity;

2. The n-sphere S^n, which is a submanifold of $\mathbb{R}^{n+1}$.

3. Lie groups (where being a smooth manifold is part of the definition);

4. The product of two manifolds [?].

1.3 Calculus on Manifolds

Within the framework of smooth structures it is possible to define concepts from analysis on a manifold M. It is usually convenient to describe these concepts in terms of local coordinates, however it is important to be sure that such definitions are coordinate independent. In this chapter, whenever a concept is defined in terms of a local coordinate chart, the coordinate independence is verified by showing the definition is invariant under a change of coordinates.

1.3.1 The Tangent Space

In $\mathbb{R}^n$ we are used to thinking of a vector as pointing from one point in space to another. In an arbitrary space this idea no longer makes sense. Instead, we must regard a vector as an object associated with a single point on the space. To develop these ideas we think about a vector as acting on functions on M.

A function $f : M \to \mathbb{R}$ is *smooth* at a point $p \in M$ if there is a coordinate chart (U, x) covering p such that the function $\bar{f} = f \circ x^{-1} : x(U) \to \mathbb{R}$ is smooth at $x(p)$. Note that this definition does not depend on the choice of chart. For suppose $f : M \to \mathbb{R}$ is smooth at p for a coordinate chart (U, x) and let (V, y) be another coordinate chart at p. Then, as $x \circ y^{-1}$ is smooth by definition, the function $f \circ y^{-1} = \bar{f} \circ x \circ y^{-1}$ is smooth.

Let $C^\infty(U)$ denote the space of smooth functions $f : U \to \mathbb{R}$. A *vector field* on U is a linear functional $X : C^\infty(U) \to C^\infty(U)$ that satisfies the *Leibniz Rule*,

$$X(fh) = fX(h) + X(f)h,$$

for $f, h \in C^\infty(U)$. Given a coordinate chart $(x^1, \ldots, x^n)$ at p, a vector field X in a neighbourhood

U of p can be written

$$X = X^i(x^1, \ldots, x^n)\frac{\partial}{\partial x^i}, \tag{1.1}$$

where the X^i are functions on U and are referred to as the *components* of X with respect to the coordinate system. The vector field is *smooth* if all of the X^i are smooth. Thus, we are able to extend the definition of a vector field to a global vector field $X : C^\infty(M) \to C^\infty(M)$. The space of all smooth vector fields on M is denoted $\mathcal{T}(M)$. The expression $X^i(p)\frac{\partial}{\partial x^i} = X^i_p\frac{\partial}{\partial x^i}$ is referred to as a *tangent vector* at p.

In equation (1.1) and here on the *Einstein Summation Convention* will apply: when an index appears twice in an expression, summation over that index is implicitly assumed. That is, the expression (1.1), should be interpreted as

$$X = \sum_i X^i \frac{\partial}{\partial x^i}.$$

Let $(x^{1'}, \ldots, x^{n'})$ be another coordinate chart at p. Using the chain rule we obtain the *contravariant transformation law*

$$X = X^i \frac{\partial}{\partial x^i} = X^i \frac{\partial x^{j'}}{\partial x^i}\frac{\partial}{\partial x^{j'}}. \tag{1.2}$$

By defining $X^{j'} = X^i\frac{\partial x^{j'}}{\partial x^i}$ we see that in this alternative coordinate chart the vector field has the same form.

The space of all tangent vectors $X^i_p\frac{\partial}{\partial x^i}$ at a point p is called the *tangent space to M at p*, denoted T_pM. It is clear that T_pM is a vector space and that for a given coordinate chart $(x^1, \ldots, x^n)$ it has basis $\frac{\partial}{\partial x^i}$. For brevity, when the coordinate chart is clear, we sometimes use the notation $\partial_i \equiv \frac{\partial}{\partial x^i}$. The disjoint union of the tangent spaces is denoted TM and is called the *tangent bundle*.

A *smooth curve* on M is a smooth map $\mathcal{C} : I \to M$ for an open interval $I = (t_1, t_2) \subset \mathbb{R}$. The curve passes through a point $p \in M$ if there is $t_p \in I$ so that $\mathcal{C}(t_p) = p$. Under a chart (U, x), the coordinate representation of $\mathcal{C}$ is denoted $\bar{\mathcal{C}} = (x^1(t), \ldots, x^n(t))$.

Let $\mathcal{C} : I \to M$ be a smooth curve with $\mathcal{C}(t_p) = p$, and $f \in C^\infty(M)$, with coordinate representations $\bar{\mathcal{C}}$ and $\bar{f}$ respectively. The composition $f \circ \mathcal{C}$ maps the interval I to $\mathbb{R}$. The *derivative, df* of f along $\mathcal{C}$ at p is

$$
\begin{aligned}
\frac{d}{dt} f \circ \mathcal{C} \bigg|_{t_p} &= \frac{d}{dt} \bar{f}(x^1(t), \ldots, x^n(t)) \\
&= \frac{\partial \bar{f}}{\partial x^i} \bigg|_{x_p} \frac{\partial x^i}{\partial t} \bigg|_{t_p} \\
&= X_p(f).
\end{aligned}
\tag{1.3}
$$

where $X_p = \frac{\partial x^i}{\partial t}\big|_{t_p} \frac{\partial}{\partial x^i}$ is the *tangent vector* to $\mathcal{C}$ at p.

1.3.2   The Cotangent Space

Consider a vector space V with basis $\{e_1, \ldots, e_n\}$ so that any vector $v \in V$ can be written $v = v^i e_i$ with respect to that basis. The *dual space* V^* to V is the vector space of linear functionals $\omega : V \to \mathbb{R}$. The basis for V^* is defined to be $\{\varepsilon^1, \ldots, \varepsilon^n\}$ such that

$$
\langle \varepsilon^i, e_j \rangle \equiv \varepsilon^i e_j = \delta^i_j,
$$

where the symbol $\langle \omega, v \rangle$ denotes the *inner product* of $v \in V$ and $\omega \in V^*$, and δ^i_j is the *Kronecker delta* defined by

$$
\delta^i_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}
$$

The dual space to the tangent space $T_p M$ is called the *cotangent space*, denoted $T_p^* M$. The elements of $T_p^* M$ are called *covectors* or *1-forms*. For a function $f \in C^\infty(M)$, we define its *exterior derivative*, df, by the unique 1-form such that

$$
df(X) = \langle df, X \rangle := X(f)
\tag{1.4}
$$

for all $X \in TM$. Observe that if we apply this definition to each of the simple functions $f = x^1, \ldots x^n$, where the basis for $T_p M$ is $(\partial_1, \ldots, \partial_n)$, we obtain

$$
\langle dx^i, \partial_j \rangle = \delta^i_j.
$$

Thus, the basis for the cotangent space is $dx^1, \ldots, mbox dx^n$.

The form for a general covector is $\omega = \omega_i dx^i$. As in the case for vectors, we require that covectors are invariant under coordinate transformations, and so can derive the *covariant transformation property*:

$$\omega = \omega_i dx^i = \omega_i \frac{\partial x^i}{\partial x^{j'}} dx^{j'}. \tag{1.5}$$

The components of ω in the new frame are $\omega_{j'} = \omega_i \frac{\partial x^i}{\partial x^{j'}}$.

1.3.3 Tensors

A (r, s)-*tensor* (say "r times contravariant and s times covariant") over a vector space V is a multilinear map

$$T : \underbrace{V^* \times \cdots \times V^*}_{r \text{ copies}} \times \underbrace{V \times \cdots \times V}_{s \text{ copies}} \to \mathbb{R}.$$

The specification that the map is multilinear means that it is linear in each argument. The rank of a tensor is $(r + s)$, that is, the total number of arguments it takes. The space of (r, s)-tensors on a vector space V is denoted by $T^r_s(V)$.

Given tensors $T \in T^r_s$ and $S \in T^p_q$, the *tensor product*, $T \otimes S \in T^{r+p}_{s+q}$ is defined by

$$T \otimes S(\omega^1, \ldots, \omega^{r+p}; v_1, \ldots, v_{s+q}) = T(\omega^1, \ldots, \omega^r; v_1, \ldots, v_s) S(\omega^{r+1}, \ldots, \omega^{r+p}; v_{s+1}, \ldots, v_{s+q})$$

The tensor product is not commutative, but it is associative, that is

$$(T \otimes S) \otimes R = T \otimes (S \otimes R),$$

so a product of three tensors can be written $T \otimes S \otimes R$ without ambiguity.

On the tangent space T_p, a (r, s)-tensor maps an ordered set of r 1-forms and s vectors to $\mathbb{R}$. We can immediately see that vectors are type $(1,0)$-tensors and 1-forms are type $(0,1)$-tensors. Scalar functions are tensors of type $(0,0)$. The components of a (r, s)-tensor over $T_p M$ can be determined in coordinates by acting on the basis vectors and 1-forms,

$$T^{i_1 \ldots i_r}_{j_1 \ldots j_s} = T(dx^{i_1}, \ldots, dx^{i_r}; \partial_{j_1}, \ldots, \partial_{j_s}),$$

where each of the i's and j's run through $1 \ldots n$. Thus, the tensor can be written in expanded form

$$T = T^{i_1 \ldots i_r}_{j_1 \ldots j_s} \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s}.$$

Under a coordinate change the *tensor transformation law* is

$$T^{i'_1 \ldots i'_r}_{j'_1 \ldots j'_s} = \frac{\partial x^{i'_1}}{\partial x^{i_1}} \cdots \frac{\partial x^{i'_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial x^{j'_1}} \cdots \frac{\partial x^{j_s}}{\partial x^{j'_s}} T^{i_1 \ldots i_r}_{j_1 \ldots j_s}. \tag{1.6}$$

This is consistent with the transformation laws already seen for vectors (equation (1.2)) and 1-forms (equation (1.5)).

The operation *contraction* is a map performed by summing over one upper and one lower index. Precisely which indices are involved can be conveniently indicated by using repeated indices in local coordinates. For example, on an n-dimensional vector space, contraction over the second upper index and first lower index on a $(3, 2)$-tensor produces:

$$T^{ijk}_{jl} = \sum_{j=1}^{n} T^{ijk}_{jl} = S^{ik}_{l}.$$

Lemma 1.3.1 *Contraction of a tensor is a tensor.*

Proof: In order to prove this we must prove that the resultant tensor obeys the transformation law (1.6). We prove the lemma for the tensor in the above example for simplicity. A proof for a general tensor proceeds in the same way.
For the tensor $T = T^{ijk}_{jl}$, (1.6) gives

$$T^{i'j'k'}_{j'l'} = \frac{\partial x^{i'}}{\partial x^{i}} \frac{\partial x^{j'}}{\partial x^{j}} \frac{\partial x^{k'}}{\partial x^{k}} \frac{\partial x^{j}}{\partial x^{j'}} \frac{\partial x^{l}}{\partial x^{l'}} T^{ijk}_{jl}.$$

Here we recall that

$$\frac{\partial x^{j}}{\partial x^{n}} \frac{\partial x^{m}}{\partial x^{j}} = \delta^{m}_{n}$$

9

where δ_n^m is the Kronecker delta. Thus we have

$$T_{j'l'}^{i'j'k'} = \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{k'}}{\partial x^k} \frac{\partial x^l}{\partial x^{l'}} T_{jl}^{ijk}.$$

That is

$$S_{l'}^{i'k'} = \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{k'}}{\partial x^k} \frac{\partial x^l}{\partial x^{l'}} S_l^{ik},$$

i.e. S_l^{ik} obeys the tensor transformation law as required. ∎

Thus, contraction is a mapping from $(r+1, s+1)$-tensors to (r, s)-tensors.

A tensor is *symmetric* in some of its indices if it is unchanged by an interchange of those indices. For example, if

$$S_{ijkl} = S_{jikl},$$

we say S_{ijkl} is symmetric in i and j. We can *symmetrize* a tensor over any number of its indices by summing over the permutations of those indices and dividing by the number of permutations. The resulting tensor will be symmetric. Symmetrization is denoted by round brackets and is given by:

$$T_{(i_1,\ldots,i_n)i_{n+1}\ldots i_k} = \frac{1}{n!} \sum_{\sigma \in S_n} T_{\sigma(i_1,\ldots,i_n)i_{n+1}\ldots i_k}, \tag{1.7}$$

where S_n is the group of permutations on $(1,\ldots,n)$.

A tensor is *skew-symmetric* in some of its indices if it changes sign under a swap of any two of those indices, for example, if

$$S_{ijkl} = -S_{ikjl},$$

we say S_{ijkl} is skew-symmetric in i and j. In a similar fashion to symmetrization, we can *skew* a tensor over a number of its indices by taking the alternating sum over the permutations of those indices and dividing by the number of permutations. This is denoted by square brackets:

$$T_{[i_1,\ldots,i_n]i_{n+1}\ldots i_k} = \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}\sigma \, T_{\sigma(i_1,\ldots,i_n)i_{n+1}\ldots i_k}, \tag{1.8}$$

where $\text{sgn}\sigma = -1$ if the permutation σ results from an odd number of swaps of indices, and $\text{sgn}\sigma = 1$ if the number of swaps is even. We can also symmetrize or skew over upper indices.

Sometimes we may wish to symmetrize or skew over indices which are not adjacent. If so, the indices not included are placed within vertical bars, for example,

$$T_{(i|j|k)} = \frac{1}{2}(T_{ijk} + T_{kji}).$$

1.3.4 Riemannian Manifolds

A *Riemannian metric* on a smooth manifold M is a symmetric, positive definite, $(0,2)$-tensor field g. That is, $g(X, X) > 0$ if $X \in T_pM \neq 0$. A *Riemannian manifold* is a smooth manifold equipped with a Riemannian metric. In coordinates, the metric has the form

$$g = g_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j.$$

For the sake of brevity the tensor product symbol is usually dropped and this is written $g = g_{ij}\mathrm{d}x^i\mathrm{d}x^j$.

On each tangent space T_pM the Riemannian metric determines an inner product,

$$\langle X_p, Y_p \rangle = g_{ij}X_p^i Y_p^j, \tag{1.9}$$

for $X_p, Y_p \in T_pM$. In particular, $\langle \partial_i, \partial_j \rangle = g_{ij}$. For example, in Euclidean space, $\mathbb{R}^n$, the metric is $g_{ij} = \delta_{ij}$. Let $X_0 = (X_0^1, \ldots, X_0^n)$ and $Y_0 = (Y_0^1, \ldots, Y_0^n)$ be two vectors at a point $p_0 \in \mathbb{R}^n$ in Cartesian coordinates. Then from (1.9) we have

$$\langle X_0, Y_0 \rangle = \delta_{ij}X_0^i Y_0^j = X_0^1 Y_0^1 + \cdots + X_0^n Y_0^n,$$

which is the familiar dot product.

Theorem 1.3.1 *Every smooth manifold M may be equipped with a Riemannian metric.*

Proof: Let $(x_i, U_i)_{i \in I}$ be an atlas on M, and let $(\psi_i)_{i \in I}$ be a partition of unity subordinate to this atlas. For $p \in U_i$ let $X_p, Y_p \in T_pM$, be given in coordinates by $(X_{pi}^1, \ldots, X_{pi}^n)$ and $(Y_{pi}^1, \ldots, Y_{pi}^n)$ respectively. Then we can define a Riemannian metric on M by

$$g(X_p, Y_p) = \sum_{i \in I} \psi_i X_{pi}^j Y_{pi}^j.$$

∎

We define the *inverse metric tensor* g^{ij} by

$$g^{ij}g_{jk} = g_{jk}g^{ij} = \delta^i_k.$$

Clearly this is also symmetric. The metric and inverse metric can be used to perform the important operations of *raising and lowering indices* on a tensor, which allows us to identify the contravariant and covariant components. In particular, we can use the metric and inverse to change vectors to covectors and vice versa:

$$\begin{aligned} X_i &= g_{ij}X^j \\ \omega^i &= g^{ij}\omega_j \end{aligned}.$$

If we take a n-submanifold M of $\mathbb{R}^m$ then our knowledge of the metric in $\mathbb{R}^m$ allows us to construct a metric on M. Let $(u^1, \ldots, u^m)$ be the usual Cartesian coordinates in $\mathbb{R}^m$ and let M have local coordinates $(x^1, \ldots, x^n)$. The *induced metric* is [?]

$$g = \sum_{i=1}^m (\mathrm{d}u^i)^2 = \sum_{i=1}^m \left(\frac{\partial u^i}{\partial x^j}\mathrm{d}x^j\right)^2. \tag{1.10}$$

1.4   Mappings between Manifolds

Consider two manifolds M and N, and suppose there exists a diffeomorphism $\phi : M \to N$. Let f be a function $f : N \to \mathbb{R}$. We can precompose the diffeomorphism ϕ with f to define a corresponding function on M. This is the *pull back* of f by ϕ:

$$\phi^* f = (f \circ \phi). \tag{1.11}$$

Recalling that a vector field on a manifold is a differential operator on the space of smooth functions on the manifold, we can use ϕ to define vector fields on N in terms of vector fields on M. This is called the *push forward* of a vector field X by ϕ, denoted $\phi_* X$, and is defined by its action on functions $f \in C^\infty(N)$,

$$(\phi_* X)f = X(\phi^* f). \tag{1.12}$$

That is, the action of the vector field $\phi_* X \in \mathcal{T}(N)$ is the action of $X \in \mathcal{T}(M)$ on the pulled back function $\phi^* f$.

The pull back and push forward can be extended to purely contravariant or purely covariant tensors. For a diffeomorphism $\phi : M \to N$, with coordinate representation $\phi(x^1, \ldots x^n) = (\phi^1(x), \ldots, \phi^m(x))$, the general formulae are [?]:

$$(\phi^* T)_{i_1, \ldots, i_s} = \frac{\partial \phi^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial \phi^{j_s}}{\partial x^{i_s}} T_{j_1 \ldots j_s} \tag{1.13}$$

$$(\phi_* T)^{j_1, \ldots, j_r} = \frac{\partial \phi^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial \phi^{j_r}}{\partial x^{i_r}} T^{i_1 \ldots i_r} \tag{1.14}$$

If $\phi : M \to N$ is a diffeomorphism and Y is a vector field on N, we put

$$\phi^* Y = (\phi^{-1})_* Y. \tag{1.15}$$

1.5 Isometries

Intuitively, we think of a symmetry on a manifold as a transformation of the manifold under which the geometry "looks the same", that is, the metric remains the same under the transformation. Mathematically, symmetries on a Riemannian manifold M with metric g are given by *isometries*; diffeomorphisms $\phi : M \to M$ for which $\phi^* g = g$. A composition of isometries is clearly also an isometry, as is the inverse of an isometry. Thus, the set of isometries on M forms a group, called the *isometry group*, $\mathcal{I}(M)$, of M. In fact, the isometry group can be shown to be a finite dimensional Lie group acting on M [?].

An important example is the Euclidean space $\mathbb{R}^n$ with metric $g_{ij} = \delta_{ij}$. By elementary geometry, the isometry group $\mathcal{I}(\mathbb{R}^n)$ turns out to be the *Euclidean group* $E(n) = \mathcal{O}(n) \times R^n$, where $\mathcal{O}(n)$ is the group of orthogonal matrices (invertible $n \times n$ matrices A such that $A^T A = \mathbb{I}$). The group acts on $x \in \mathbb{R}^n$ by

$$x \mapsto A(x) + b$$

where $b \in \mathbb{R}^n$ is a translation, and $A \in \mathcal{O}(n)$ corresponds to orthogonal transformations, that is, the reflections and rotations. The connected component of $E(n)$ containing the identity forms the

continuous subgroup of *orientation preserving* isometries. The structure $SE(n) = \mathcal{SO}(n) \times \mathbb{R}^n$, where $\mathcal{SO}(n) = \{A \in \mathcal{O}(n) |\det A = 1\}$ corresponds to the rotations. The reflections are *orientation reversing.*

A *1-parameter family of diffeomorphisms* is a group of diffeomorphisms $\phi_t : M \to M$, $t \in \mathbb{R}$ such that $\phi_t \circ \phi_s = \phi_{t+s}$. The identity is ϕ_0 and $\phi_{-t} = \phi_t^{-1}$. On an n-manifold M there is a 1-1 correspondence between 1-parameter families of diffeomorphisms and vector fields [?]. Under a family of diffeomorphisms ϕ_t, each point p traces out a smooth curve $\mathcal{C}_p(t) = (x_p^1(t), \ldots, x_p^n(t))$ with $\mathcal{C}(0) = p$. We can define a vector field $X^i \frac{\partial}{\partial x^i}$ along $\mathcal{C}_p$ (cf. (1.3)) by

$$\frac{dx_p^i}{dt} = X^i. \tag{1.16}$$

This can be done for every point on our manifold and thus we have a vector field on M. The curves $\mathcal{C}$ are called *integral curves* of X. Now, we wish to view this as the diffeomorphisms parameterised by t acting on a point $p \in M$, rather than a curve parameterised by t through the point p, so we write

$$\phi_t(p) = \mathcal{C}_p(t).$$

The vector field X is called the *generator* of the family of diffeomorphisms.

The diffeomorphisms R_θ clearly form a 1-parameter family with parameter θ. The corresponding vector field is

$$X = -x^2 \frac{\partial}{\partial x^1} + x^1 \frac{\partial}{\partial x^2}. \tag{1.17}$$

The integral curves of X are concentric circles about the origin.

1.6   The Lie Derivative

On a manifold M we may ask how a tensor T changes as it travels along the integral curves of a vector field. To this end we define the *Lie Derivative*, $\mathcal{L}_X T$, of a tensor field $T \in T_s^r$ along a vector field X by

$$\mathcal{L}_X T := \frac{d}{dt}(\phi_t^* T)|_{t=0} \tag{1.18}$$

where ϕ_t is the 1-parameter family of diffeomorphisms generated by X. The Lie derivative is a well-behaved differential operator, that is, for real numbers a, b and tensors T, S,

1. it is linear:
$$\mathcal{L}_X(aT + bS) = a\mathcal{L}_X T + b\mathcal{L}_X S \qquad (1.19)$$

2. it obeys the Leibniz Rule:
$$\mathcal{L}_X(T \otimes S) = (\mathcal{L}_X T) \otimes S + T \otimes (\mathcal{L}_X S) \qquad (1.20)$$

We use these properties to give explicit formulae for functions, 1-forms and vector fields, then extend them to a general tensor. In the working that follows, M is a Riemannian n-manifold with local coordinates $(x^1, \ldots, x^n)$ in which the vector field is $X = X^i \frac{\partial}{\partial x^i}$.

Theorem 1.6.1 *Let $f : M \to \mathbb{R}$ be a smooth function on M. Then the Lie derivative of f with respect to X is*
$$\mathcal{L}_X(f) = Xf, \qquad (1.21)$$
where the right hand side is the action of the vector field on the function as in §1.3.1.

Proof:

$$
\begin{aligned}
\mathcal{L}_X(f) &= \frac{d}{dt}(\phi_t^* f)|_{t=0} \\
&= \frac{d}{dt}(f \circ \phi_t)|_{t=0} \\
&= X^i \frac{\partial f}{\partial x^i} \qquad \text{(cf. equation (1.3))} \\
&= X(f)
\end{aligned}
$$

$\blacksquare$

Unlike functions, we are unable to write the Lie derivative of a 1-form in a manifestly coordinate free manner. Thus, we must operate in a local chart and check the coordinate independence explicitly.

Theorem 1.6.2 *Let $\omega = \omega_i dx^i$ be a 1-form with respect to the coordinate system $(x^1, \ldots, x^n)$. The Lie derivative of ω with respect to X is*

$$\mathcal{L}_X \omega = \left(\omega_i \frac{\partial X^i}{\partial x^j} + X^i \frac{\partial \omega_j}{\partial x^i} \right) dx^j. \tag{1.22}$$

Proof: From (1.13) we have

$$\phi^* \omega = \omega_i \frac{\partial \phi^i}{\partial x^j} dx^j.$$

Consider the action of the Lie derivative on the basis 1-forms dx^i:

$$
\begin{aligned}
\frac{d}{dt} (\phi_t^* dx^i)|_{t=0} &= \frac{d}{dt} \left(\frac{\partial \phi_t^i}{\partial x^j} dx^j \right) |_{t=0} \\
&= \frac{d}{dt} \frac{\partial \phi_t^i}{\partial x^i} \bigg|_{t=0} dx^j \\
&= \frac{\partial X^i}{\partial x^j} dx^j \qquad (\text{since } \phi_0 = id).
\end{aligned} \tag{1.23}
$$

Now consider the general 1-form $\omega_i dx^i$. Applying the Leibniz rule,

$$
\begin{aligned}
\mathcal{L}_X(\omega_i dx^i) &= \omega_i \mathcal{L}_X(dx^i) + \mathcal{L}_X(\omega_i) dx^i \\
&= \omega_i \frac{\partial X^i}{\partial x^j} dx^j + X^j \frac{\partial \omega_i}{\partial x^j} dx^i \\
&= \left(\omega_i \frac{\partial X^i}{\partial x^j} + X^i \frac{\partial \omega_j}{\partial x^i} \right) dx^j,
\end{aligned}
$$

where in the last term we have swapped dummy indices. ∎

Now we must confirm the coordinate independence of this result. Let x' be a change of coordinates. Then by equation (1.2) in this coordinate system

$$X = X^{j'} \frac{\partial}{\partial x^{j'}} = X^i \frac{\partial x^{j'}}{\partial x^i} \frac{\partial}{\partial x^{j'}},$$

and by equation (1.5),

$$\omega = \omega_{j'} dx^{j'} = \omega_i \frac{\partial x^i}{\partial x^{j'}} dx^{j'}.$$

Substituting this into (1.22) in the x' frame:

$$
\begin{aligned}
\mathcal{L}_X\omega &= \left(X^{i'}\frac{\partial}{\partial x^{i'}}\omega_{j'} + \omega_{i'}\frac{\partial}{\partial x^{j'}}X^{i'} \right)dx^{j'}\\[4pt]
&= \left(X^m\frac{\partial x^{i'}}{\partial x^m}\frac{\partial}{\partial x^{i'}}\left(\omega_n\frac{\partial x^n}{\partial x^{j'}} \right) + \omega_n\frac{\partial x^n}{\partial x^{i'}}\frac{\partial}{\partial x^{j'}}\left(X^m\frac{\partial x^{i'}}{\partial x^m} \right) \right)dx^{j'}\\[4pt]
&= \left(X^m\frac{\partial x^{i'}}{\partial x^m}\left(\frac{\partial \omega_n}{\partial x^{i'}}\frac{\partial x^n}{\partial x^{j'}} + \omega_n\frac{\partial^2 x^n}{\partial x^{i'}\partial x^{j'}} \right) + \omega_n\frac{\partial x^n}{\partial x^{i'}}\left(\frac{\partial X^m}{\partial x^{j'}}\frac{\partial x^{i'}}{\partial x^m} + X^m\frac{\partial^2 x^{i'}}{\partial x^{i'}\partial x^m} \right) \right)dx^{i'}\\[4pt]
&= \left(X^m\frac{\partial x^{i'}}{\partial x^m}\frac{\partial \omega_n}{\partial x^{i'}}\frac{\partial x^n}{\partial x^{j'}} + X^m\omega_n\frac{\partial x^{i'}}{\partial x^m}\frac{\partial^2 x^n}{\partial x^{i'}\partial x^{j'}} + \omega_n\frac{\partial x^n}{\partial x^{i'}}\frac{\partial X^m}{\partial x^{j'}}\frac{\partial x^{i'}}{\partial x^m} + \omega_n X^m\frac{\partial x^n}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^{j'}\partial x^m} \right)dx^{j'}\\[4pt]
&= \left(\left(X^m\frac{\partial \omega_n}{\partial x^{j'}} + \omega_n\frac{\partial X^m}{\partial x^{j'}} \right)\delta^n_m + \omega_n X^m\left(\frac{\partial x^{i'}}{\partial x^m}\frac{\partial^2 x^n}{\partial x^{i'}\partial x^{j'}} + \frac{\partial x^n}{\partial x^{i'}}\frac{\partial^2 x^{i'}}{\partial x^{j'}\partial x^m} \right) \right)dx^{j'}. \qquad (1.24)
\end{aligned}
$$

Here we notice that

$$
\frac{\partial y^i}{\partial x^m}\frac{\partial x^n}{\partial y^i} = \delta^n_m,
$$

therefore

$$
\frac{\partial}{\partial y^j}\left(\frac{\partial y^i}{\partial x^m}\frac{\partial X^n}{\partial y^i} \right) = \frac{\partial^2 y^i}{\partial y^j\partial x^m}\frac{\partial x^n}{\partial y^i} + \frac{\partial y^i}{\partial x^m}\frac{\partial^2 x^n}{\partial y^j\partial y^i} = 0.
$$

Thus the second term in the expression (1.24) vanishes and we have

$$
\mathcal{L}_X\omega = \left(X^m\frac{\partial x^{j'}}{\partial x^n}\frac{\partial \omega_m}{\partial x^{j'}} + \omega_m\frac{\partial x^{j'}}{\partial x^n}\frac{\partial X^m}{\partial x^{j'}} \right)\frac{\partial x^n}{\partial x^{j'}}dx^{j'} = \left(X^m\frac{\partial}{\partial x^m}\omega_n + \omega_m\frac{\partial}{\partial x^n}X^m \right)dx^n.
$$

Hence proving coordinate independence.

Theorem 1.6.3 *Let Y be a vector field. The Lie derivative of Y with respect to X is*

$$
\mathcal{L}_X Y = [X, Y], \qquad (1.25)
$$

where $[\,,\,]$ denotes the Lie bracket:

$$
[X, Y] = XY - YX.
$$

While this is inherently coordinate independent, the proof requires working in coordinates.

Proof: Consider the left hand side first. Let $Y = \frac{\partial}{\partial x^j}$ and recall $\langle \frac{\partial}{\partial x^j}, \mathrm{d}x^i \rangle = \delta_j^i$. It follows that

$$\mathcal{L}_X \langle \frac{\partial}{\partial x^j}, \mathrm{d}x^i \rangle = \left(\mathcal{L}_X \frac{\partial}{\partial x^j} \right) \mathrm{d}x^i + \left(\mathcal{L}_X \mathrm{d}x^i \right) \frac{\partial}{\partial x^j} = 0,$$

therefore, from (1.23) we have

$$\mathcal{L}_X \frac{\partial}{\partial x^j} = -\frac{\partial X^i}{\partial x^j} \frac{\partial}{\partial x^i}.$$

Now consider $Y = Y^j \frac{\partial}{\partial x^j}$

$$
\begin{aligned}
\mathcal{L}_X (Y^j \frac{\partial}{\partial x^j}) &= \mathcal{L}_X(Y^j) \frac{\partial}{\partial x^j} + Y^j \mathcal{L}(\frac{\partial}{\partial x^j}) \\
&= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j} - Y^i \frac{\partial X^i}{\partial x^j} \frac{\partial}{\partial x^i} \\
&= \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}
\end{aligned}
$$

(again, swapping dummy indices in the last step).

Now consider the right hand side of (1.25). The action of a product of vector fields on a function $(XY)f$ is $X(Yf)$ so we have

$$
\begin{aligned}
(XY - YX)f &= X^i \frac{\partial}{\partial x^i} \left(Y^j \frac{\partial f}{\partial x^j} \right) - Y^i \frac{\partial}{\partial x^i} \left(X^j \frac{\partial f}{\partial x^j} \right) \\
&= X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial f}{\partial x^j} + X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j} - Y^i \frac{\partial X^j}{\partial x^i} \frac{\partial f}{\partial x^i} - Y^i X^j \frac{\partial^2 f}{\partial x^i \partial x^j} \\
&= \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial f}{\partial x^j},
\end{aligned}
$$

that is, the same as the left hand side. Thus the result is proved. ∎

Now we can deduce the Lie derivative of a general tensor. Consider the arbitrary (r, s)-tensor

$$T = T^{i_1 \dots i_r}_{j_1 \dots j_s} \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s}$$

By the linearity (1.19) and the Leibniz rule (1.20), we find Lie derivative of T to be

$$
\begin{aligned}
\mathcal{L}_X T \; = \; & (X T^{i_1 \dots i_r}_{j_1 \dots j_s}) \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s} \\
& + T^{i_1 \dots i_r}_{j_1 \dots j_s} \left[\sum_{n=1}^{r} \partial_{i_1} \otimes \cdots \otimes (\mathcal{L}_X \partial_{i_n}) \otimes \dots \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s} \right. \\
& \left. + \sum_{m=1}^{s} \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes (\mathcal{L}_X \mathrm{d}x^{j_m}) \otimes \cdots \otimes \mathrm{d}x^{j_s} \right].
\end{aligned}
$$

This is coordinate independent by virtue of the independence of the tensor product and since we have already shown independence for (1,0)- and (0,1)-tensors.

Of particular interest is the covariant 2-tensor, that is tensors of the same type as the Riemannian metric. The Lie derivative for a covariant 2-tensor $\varphi = \varphi_{ij} dx^i \otimes dx^j$ is:

$$
\mathcal{L}_X \varphi = \left(X^i \frac{\partial}{\partial x^i} \varphi_{jk} + \frac{\partial X^i}{\partial x^j} \varphi_{ik} + \frac{\partial X^i}{\partial x^k} \varphi_{ji} \right) dx^j \otimes dx^k \tag{1.26}
$$

1.7   The Covariant Derivative

A *linear connection* on M is a map

$$
\nabla : \mathcal{T}(M) \times \mathcal{T}(M) \to \mathcal{T}(M)
$$

written $\nabla_X Y$, such that

1. $\nabla_X Y$ is linear over $\mathbb{R}$ in Y:

$$
\nabla_X (aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2,
$$

2. $\nabla_X Y$ is linear over f in X:

$$
\nabla_{(fX_1 + gX_2)} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y,
$$

3. and ∇ obeys the Leibniz rule:

$$
\nabla_X (fY) = f\nabla_X Y + (Xf)Y.
$$

$\nabla_X Y$ is the *covariant derivative* of Y in the direction X.

For a given vector field X, the covariant derivative is a map from the space of vector fields into itself. With respect to a local coordinate system (x, U), we can consider the covariant derivative of the basis vectors along each coordinate direction. This will define a $n \times n$ matrix of components for the resulting vector field:

$$\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k.$$

The Γ_{ik}^j are called the *Christoffel Symbols*. We write $\nabla_{\partial_i} \equiv \nabla_i$. As the following lemma shows, for general vector fields X, Y, the action of a connection is completely determined by its Christoffel symbols [?].

Lemma 1.7.1 *Let ∇ be a linear connection on U and let $X, Y \in \mathcal{T}(U)$ be written $X = X^i \partial_i$ and $Y = Y^j \partial_j$ in local coordinates. Then*

$$\nabla_X Y = (XY^k + X^i Y^j \Gamma_{ij}^k) \partial_k$$

Proof:

$$
\begin{aligned}
\nabla_X Y &= \nabla_X (Y^j \partial_j) \\
&= Y^j \nabla_X \partial_j + (XY^j) \partial_j && \text{(Leibniz)} \\
&= Y^j \nabla_{X^i \partial_i} \partial_j + (XY^j) \partial_j \\
&= X^i Y^j \nabla_{\partial_i} \partial_j + (XY^j) \partial_j && (C^\infty\text{-linearity in X)} \\
&= X^i Y^j \Gamma_{ij}^k \partial_k + (XY^j) \partial_j \\
&= (XY^k + X^i Y^j \Gamma_{ij}^k) \partial_k
\end{aligned}
$$

where in the last line we have simply changed the dummy index. ∎

In particular, the above lemma tells us

$$\nabla_i Y^j = \partial_i Y^j + \Gamma_{ik}^j Y^k. \tag{1.27}$$

We can use the above expression to derive coordinate transformation properties for the Christoffel symbols. We constructed the Christoffel symbols so that the left hand side of (1.27) transforms as a tensor. Therefore, from equation (1.6), under a coordinate change $x \to x'$ we require

$$\nabla_{i'} Y^{j'} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \nabla_i Y^j. \tag{1.28}$$

Expanding both sides of (1.28)

$$\frac{\partial}{\partial x^{i'}} Y^{j'} + \Gamma^{j'}_{i'k'} Y^{k'} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \left(\frac{\partial}{\partial x^i} Y^j + \Gamma^j_{ik} Y^k \right)$$

$$\frac{\partial x^i}{\partial x^{i'}} \frac{\partial}{\partial x^i} \left(\frac{\partial x^{j'}}{\partial x^j} Y^j \right) + \Gamma^{j'}_{i'k'} \frac{\partial x^{k'}}{\partial x^k} Y^k = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial}{\partial x^i} Y^j + \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \Gamma^j_{ik} Y^k$$

$$\frac{\partial x^i}{\partial x^{i'}} \frac{\partial^2 x^{j'}}{\partial x^i \partial x^j} Y^j + \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial}{\partial x^i} Y^j + \Gamma^{j'}_{i'k'} \frac{\partial x^{k'}}{\partial x^k} Y^k = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial}{\partial x^i} Y^j + \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \Gamma^j_{ik} Y^k$$

$$\Gamma^{j'}_{i'k'} \frac{\partial x^{k'}}{\partial x^k} Y^k = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \Gamma^j_{ik} Y^k - \frac{\partial x^i}{\partial x^{i'}} \frac{\partial^2 x^{j'}}{\partial x^i \partial x^j} Y^j$$

$$\Gamma^{j'}_{i'k'} Y^k = \frac{\partial x^k}{\partial x^{k'}} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \Gamma^j_{ik} Y^k - \frac{\partial x^k}{\partial x^{k'}} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial^2 x^{j'}}{\partial x^i \partial x^k} Y^k$$

Where in the last term the dummy index has been changed $j \to k$. The law must hold for all Y, thus the transformation law for Christoffel symbols is

$$\Gamma^{j'}_{i'k'} = \frac{\partial x^k}{\partial x^{k'}} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^{j'}}{\partial x^j} \Gamma^j_{ik} - \frac{\partial x^k}{\partial x^{k'}} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial^2 x^{j'}}{\partial x^i \partial x^k}. \tag{1.29}$$

Note that (1.29) is not the tensor transformation law. This is because the Christoffel symbols themselves are not tensors.

21

The linear connection ∇ can be extended to act over tensor fields so that the resultant connection (also denoted ∇) [?]:

1. agrees with the linear connection over the tangent space,

2. gives the ordinary derivative on functions:

$$\nabla_X f = Xf,$$

3. obeys the Leibniz rule when acting on tensor products:

$$\nabla_X(T \otimes S) = (\nabla_X T) \otimes S + T \otimes (\nabla_X S),$$

4. and commutes with contraction

$$\nabla(T^{i_1,\ldots k,\ldots,i_r}_{j_1,\ldots,k,\ldots,j_s}) = (\nabla T)^{i_1,\ldots k,\ldots,i_r}_{j_1,\ldots,k,\ldots,j_s}.$$

Lee [?] offers the following Lemma for finding the covariant derivative of a tensor field:

Lemma 1.7.2 *Let ∇ be a linear connection. The components of the covariant derivative of an (r,s)-tensor field T with respect to a coordinate system are given by*

$$\nabla_k T^{i_1\ldots i_r}_{j_1\ldots j_s} = \partial_k T^{i_1\ldots i_r}_{j_1\ldots j_s} + \sum_{l=1}^{r} T^{i_1\ldots p\ldots i_r}_{j_1\ldots j_s}\Gamma^{i_l}_{kp} - \sum_{l=1}^{s} T^{i_1\ldots i_r}_{j_1\ldots p\ldots j_s}\Gamma^{p}_{kj_l} \tag{1.30}$$

A linear connection is *metric compatible* if for all vector fields X, Y and Z the following product rule holds:

$$\nabla_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle.$$

Lemma 1.7.3 *A connection ∇ is metric compatible if and only if $\nabla g \equiv 0$*

Proof: Suppose ∇ is metric compatible.

$$\nabla_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \tag{1.31}$$

Expanding the left hand side,

$$
\begin{aligned}
\nabla_X \langle Y, Z \rangle &= \nabla_X (g_{mn} Y^m Z^n) \\
&= (\nabla_X g_{mn}) Y^m Z^n + g_{mn} (\nabla_X Y^m) Z^n + g_{mn} Y^m (\nabla_X Z^n) \\
&= (\nabla_X g_{mn}) Y^m Z^n + \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle .
\end{aligned}
$$

Therefore,

$$
(\nabla_X g_{mn}) Y^m Z^n = 0.
$$

This is true for all X, Y, Z, so

$$
\nabla g = 0
$$

$\blacksquare$

Therefore, for a metric compatible connection

$$
\begin{aligned}
\nabla_i (Y_k) &= \nabla_i (g_{jk} Y^j) \\
&= (\nabla_i g_{jk}) Y^j + (\nabla_i Y^j) g_{jk} \\
&= (\nabla_i Y^j) g_{jk}.
\end{aligned}
$$

1.8 Derivation of the Levi-Civita Connection

Consider the equation

$$
\mathcal{L}_X g = \frac{\partial}{\partial x^j}(X_k) + \frac{\partial}{\partial x^k}(X_j) - X_l g^{il} \left(\frac{\partial}{\partial x^j} g_{ik} + \frac{\partial}{\partial x^k} g_{ji} - \frac{\partial}{\partial x^i} g_{jk} \right) = 0. \tag{1.32}
$$

Define as Christoffel symbols

$$
\Gamma^l_{jk} = \frac{1}{2} g^{il} \left(\frac{\partial}{\partial x^j} g_{ik} + \frac{\partial}{\partial x^k} g_{ji} - \frac{\partial}{\partial x^i} g_{jk} \right) \tag{1.33}
$$

It is possible (but tedious) to check that these transform as Christoffel symbols should. By the symmetry of the metric tensor, these Christoffel symbols Γ^l_{jk} are symmetric in j and k. A connection has this property if and only if it is *torsion free*, that is, ∇ satisfies

$$
\nabla_X Y - \nabla_Y X = [X, Y]. \tag{1.34}
$$

23

With these Christoffel symbols, (1.32) becomes:

$$\partial_j(X_k) + \partial_k(X_j) - 2X_i\Gamma^i_{jk} = 0. \tag{1.35}$$

Now define a connection ∇ by:

$$\nabla_j X_k = \partial_j X_k - \Gamma^l_{jk} X_l. \tag{1.36}$$

Which gives

$$\mathcal{L}_X g_{jk} = 2\nabla_{(j} X_{k)} \tag{1.37}$$

This connection is metric compatible: from (1.30),

$$
\begin{aligned}
\nabla_k g_{ij} &= \partial_k g_{ij} - \Gamma^l_{ki} g_{lj} - \Gamma^l_{kj} g_{il} \\
&= \partial_k g_{ij} - \tfrac{1}{2} g^{ml} g_{lj} (\partial_k g_{im} + \partial_i g_{km} - \partial_m g_{ki}) - \tfrac{1}{2} g^{lm} g_{il} (\partial_k g_{jm} + \partial_j g_{km} - \partial_m g_{kj}) \\
&= \partial_k g_{ij} - \tfrac{1}{2}(\partial_k g_{ij} + \partial_i g_{kj} - \partial_j g_{ki}) - \tfrac{1}{2}(\partial_k g_{ji} + \partial_j g_{ki} - \partial_i g_{kj}) \\
&= 0
\end{aligned}
\tag{1.38}
$$

Theorem 1.8.1 *Fundamental Lemma of Riemannian Geometry Let (M, g) be a Riemannian Manifold. There exists a unique linear connection ∇ on M that is compatible and torsion free.*

Proof: We have already shown existence by construction, so all that is left is to show uniqueness. Following the proof in [**?**], we do this by deriving a formula for ∇. By metric compatibility (1.31) we have the following three equations:

$$
\begin{aligned}
X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle \\
Y \langle Z, X \rangle &= \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle \\
Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle
\end{aligned}
$$

Using the torsion free condition (1.34) in each of the above we get

$$
\begin{aligned}
X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_Z X \rangle + \langle Y, [X, Z] \rangle \\
Y \langle Z, X \rangle &= \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_X Y \rangle + \langle Z, [Y, X] \rangle \\
Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Y Z \rangle + \langle X, [Z, Y] \rangle.
\end{aligned}
$$

Adding the first two equations and subtracting the third

$$X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle = 2 \langle \nabla_X Y, Z \rangle + \langle Y, [X, Z] \rangle + \langle Z, [Y, X] \rangle - \langle X, [Z, Y] \rangle .$$

Now we can solve for $\langle \nabla_X Y, Z \rangle$,

$$\langle \nabla_X Y, Z \rangle = \frac{1}{2} \left(X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle - \langle Y, [X, Z] \rangle - \langle X, [Y, X] \rangle + \langle X, [Z, Y] \rangle \right) .$$

Noticing that the right hand side of the above equation does not depend on the connection, it follows that if ∇^1 and ∇^2 are two connections

$$\langle \nabla_X^1 Y, Z \rangle = \langle \nabla_X^2 Y, Z \rangle$$

for all X, Y and Z. This can only happen if $\nabla^1 = \nabla^2$. ∎

This connection is called the *metric connection* or the *Levi-Civita connection*.

1.9 Parallel Transport, geodesics and the Riemann curvature tensor

Consider a closed curve $\mathcal{C}(t) = (x^1(t), ... x^n(t))$ which passes through a point $p \in M$ at t_p, i.e. $\mathcal{C}(t_p) = p$. The tangent vector to $\mathcal{C}$ at p is given by $X_p = \frac{dx^i}{dt} |_{t_p} \partial_i = X^i \partial_i$. A vector V^j undergoes *parallel transport* along $\mathcal{C}$ if at all points on $\mathcal{C}$

$$X^i \nabla_i V^j = 0. \tag{1.39}$$

This generalises as expected to an arbitrary tensor.

Imagine you're standing at the north pole and you walk south until you reach the equator. Then you sidestep 90 degrees around the equator before walking backwards until you return to the north pole. Although on your travels you maintained the direction you were facing, you now appear to be facing at right angles to where you started. If you consider different paths for our example, such as traversing just 45 degrees around the equator instead of 90, it is clear that you would be facing at 45 degrees to your starting position. Thus the difference in angle is path

dependent. This is because of the curvature of the earth- if you traversed a closed loop on flat space you would end up facing the same way regardless of path.

By the definition of parallel transport, the covariant derivative of a tensor in a direction along which it is parallelly transported is zero. Therefore, the covariant derivative of a tensor in an arbitrary direction is a measure of the the change in the tensor compared to if it had been parallelly transported and as such is a measure of curvature. The commutator of two covariant derivatives ∇_i and ∇_j is the difference between parallelly transporting the tensor along the x^i then x^j compared to x^j then x^i.

Consider this action on a dual vector field ω_k and a smooth function f.

$$
\begin{aligned}
\nabla_i \nabla_j (f\omega_k) &= \nabla_i(\omega_k \nabla_j f + f\nabla_j \omega_k) &= \nabla_i \omega_k \nabla_j f + \omega_k \nabla_i \nabla_j f + \nabla_i f \nabla_j \omega_k + f\nabla_i \nabla_j \omega_k \\
\nabla_j \nabla_i (f\omega_k) &= \nabla_j(\omega_k \nabla_i f + f\nabla_i \omega_k) &= \nabla_j \omega_k \nabla_i f + \omega_k \nabla_j \nabla_i f + \nabla_j f \nabla_i \omega_k + f\nabla_i \nabla_j \omega_k
\end{aligned}
\tag{1.40}
$$

Subtract the second equation from the first, and using the torsion free property of the connection find that the commutator gives:

$$
\nabla_{[i}\nabla_{j]}(f\omega_k) = f(\nabla_i \nabla_j - \nabla_j \nabla_i)\omega_k
$$

Now consider a second 1-form ω'_k such that $\omega'_k = \omega_k$ at a point p. Then we can find a set of 1-forms μ_k^α and smooth functions f_α with $f_\alpha(p) = 0$ such that $\omega'_k - \omega_k = \sum_\alpha f_\alpha \mu_k^\alpha$. Therefore,

$$
(\nabla_i \nabla_j - \nabla_j \nabla_i)(\omega'_k - \omega_k) \mid_p = \sum_\alpha f_\alpha (\nabla_i \nabla_j - \nabla_j \nabla_i)\mu_k^\alpha \mid_p = 0
$$

$$
(\nabla_i \nabla_j - \nabla_j \nabla_i)\omega'_k = (\nabla_i \nabla_j - \nabla_j \nabla_i)\omega_k
$$

. Therefore, $(\nabla_i \nabla_j - \nabla_j \nabla_i)\omega_k$ can only depend on $\omega_k(p)$ and $\nabla_i \nabla_j - \nabla_j \nabla_i$ is a $(1,3)$ tensor. This defines the *Riemann Curvature Tensor*:

$$
R_{ijk}^m \omega_m = (\nabla_i \nabla_j - \nabla_j \nabla_i)\omega_k
\tag{1.41}
$$

As usual we can lower the m index on the Riemann tensor using the metric. Then it can be seen that the following symmetry properties hold [?]:

1. $R_{ijkm} = -R_{jikm}$ (skew in first two indices)

2. $R_{ijkm} = -R_{ijmk}$ (skew in last two indices)

26

3. $R_{m[kij]} = 0$ (cyclic property)

Some manipulation using properties 1 to 3 will give you another useful symmetry [?]:

4. $R_{mkij} = R_{ijmk}$ (symmetric under index pair interchange).

The Riemann tensor gives rise to two additional important tensors. The first is the *Ricci curvature tensor* which is defined to be contraction over the first and third components of the Riemann tensor

$$R_{kj} = R^i_{kij}. \tag{1.42}$$

By the symmetry property, of the Riemann curvature tensor, it is clear that the Ricci tensor is symmetric. The second is the *scalar curvature*, which is the trace of the Ricci curvature:

$$R = g^{kj}R_{kj} = R^j_j. \tag{1.43}$$

For an n-manifold, we may at first naively suppose that the Riemann tensor has n^4 independent components. However, consideration of the symmetry properties shows that not all of these are independent. As R_{mjik} is skew in the first two and last two components, this gives us $\binom{n}{2}$ independent components for each pair. The cyclic property imposes $\binom{n}{3}$ constraints. Thus, the number of independent components of the Riemann tensor is

$$\binom{n}{2}^2 - n\binom{n}{3} = \frac{n^2(n^2-1)}{12}. \tag{1.44}$$

For a surface $n = 2$, and there is one independent component. By inspection we see that $(g_{mi}g_{kj} - g_{mj}g_{ki})$ satisfies the symmetry properties, therefore for a surface the Riemann curvature tensor is

$$R_{mkij} = K(g_{mi}g_{kj} - g_{mj}g_{ki}) \tag{1.45}$$

where K is a scalar function. If we calculate the Ricci tensor we get

$$R_{kj} = Kg^{mi}(g_{mi}g_{kj} - g_{mj}g_{ki}) = Kg_{kj}, \tag{1.46}$$

and then the scalar curvature is

$$R = g^{kj}R_{kj} = Kg^{kj}g_{kj} = 2K. \tag{1.47}$$

The function K is called the *Gaussian curvature*. Note that for manifolds of dimension $n > 2$, the Gaussian curvature can still be defined, but is not necessarily scalar. Notice that

$$2\nabla_{[i}F_{j]k} = \nabla_i\nabla_j X_k - \nabla_j\nabla_i X_k, \tag{1.48}$$

therefore, by comparing (1.48) with (**??**) and (1.41), we see that the Riemann tensor can also be written

$$R^m_{kij}X_m = (\nabla_i\nabla_j - \nabla_j\nabla_i)X_k.$$

1.10   Geodesics

A *geodesic* is a path that parallel transports its own tangent vector. Expanding (1.39) gives:

$$\begin{aligned}
\frac{dx^i}{dt}\nabla_i\frac{dx^j}{dt} &= 0 \\
\frac{dx^i}{dt}\left(\partial_i\frac{dx^j}{dt} + \Gamma^j_{ik}\frac{dx^k}{dt}\right) &= 0 \\
\frac{d^2x^j}{dt^2} + \Gamma^j_{ik}\frac{dx^i}{dt}\frac{dx^k}{dt} &= 0
\end{aligned} \tag{1.49}$$

The last equation above is called the *geodesic equation.* In flat space, where we can choose coordinates such that $\Gamma^j_{ik} = 0$, the geodesic equation is simply $\frac{d^2x^j}{dt^2} = 0$, the equation for a straight line. In curved space, geodesics can be thought of as representing the shortest distance between points.

In flat space we know that parallel straight lines never meet, but obviously this is not true for geodisics in curved space. To explore this further consider a 1-parameter family of geodesics $\gamma_s(t)$, $s \in \mathbb{R}$, that do not cross. These define a smooth two dimensional surface on which we can choose the coordinates to be s and t so that the points on the surface are $x^j(s,t)$. The vector fields $T^j = \frac{\partial x^j}{\partial t}$ and $S^j = \frac{\partial x^j}{\partial s}$ represent the the tangent vectors to the geodesics and the deviation vectors respectively. The deviation vectors can be interpreted as pointing from one geodesic to the next. We want to explore how these change, so we define $V^i = (\nabla_T S)^i = T^j\nabla_j S^i$, which we interpret as the relative velocity of the geodesics, and $A^i = (\nabla_T V)^i = T^j\nabla_j V^i$ the relative acceleraton between geodesics. It can be shown (see, eg. [?]) that the relative acceleration between geodesics is proportional to the curvature:

$$A^i = R^i_{jkl}T^jT^k S^l \tag{1.50}$$

28

This is known as the *geodesic deviation equation*.

1.11   Euclidean and Hyperbolic space

Consider the following family of metrics parameterised by p:

$$g = \frac{1}{(1 + p((x^1)^2 + (x^2)^2))^2}((\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2).$$

When $p = 1$, this is the (now familiar) metric for the sphere S^2. When $p = 0$, g is the metric for Euclidean space $\mathbb{R}^2$, and when $p = -1$, and we confine our attention to the unit disc $\|x\| < 1$, we have *hyperbolic space* $\mathbb{H}^2$. These three spaces are are the 2-dimensional versions of the "model spaces" for Riemannian geometry.

The Riemann curvature for each of these spaces satisfies (1.45). As done for the sphere in the previous section, we find the Gaussian curvature by calculating one component of the tensor using the definition (1.41):

$$R^1_{212} = \frac{4p}{(1 + p((x^1)^2 + (x^2)^2))^2}.$$

Using (1.45), we find the Gaussian curvature $K = 2p$. Thus, each of these spaces have constant Gaussian curvature, and therefore are maximally symmetric.